2017

Decision making under uncertainties for renewable energy and precision agriculture

Qi Li
Iowa State University

Decision making under uncertainties for renewable energy and precision agriculture

by

Qi Li

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-majors: Industrial and Manufacturing Systems Engineering; Statistics

Program of Study Committee:
Guiping Hu, Co-Major Professor
Huaiqing Wu, Co-Major Professor
Baskar Ganapathysubramanian
Mingyi Hong
Daniel J Nordman
Sarah M Ryan

The student author and the program of study committee are solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2017

DEDICATION

In dedication to my wife for supporting me all the way!

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to many people who made this dissertation possible and an unforgettable experience for me.

First of all, I would like to express my sincerest gratitude to my advisor, Dr. Guiping Hu and Dr. Huaiqing Wu, for their continuous advice and encouragement throughout the course of my Co-PhD program and this dissertation. I would like to express my very sincere sense to Dr. Baskar Ganapathysubramanian, Dr. Mingyi Hong, Dr. Daniel J. Nordman, and Dr. Sarah M. Ryan to serve as my committee members and give me their guidance and support all the way.

In addition, I want to also offer my appreciation to my friends, colleagues, all course instructors, and department faculty and staff for making my life at Iowa State University an enjoyable experience.

Finally, thanks to my family for their unconditional patience, encouragement, care and love.

ABSTRACT

In this dissertation, mathematical programming models and statistical analysis tools have been formulated and designed to study the strategic and optimal solutions to allocate the resources and manage the risk for the renewable energy and precision agriculture. The dissertation, which consists of four papers, lies at the interface of optimization, simulation, and statistical analysis, with a focus on decision making under uncertainty for biofuel process design, renewable energy supply chain management and precision agriculture.

Bio-oil gasification which integrates fast pyrolysis and gasification processes is a relative new conversion technology and this integrated biofuel production pathway has been promoted to take advantage of economies of scale and logistic efficiency. The design of the supply chain networks, especially under uncertainties, is one of the most important decisions faced by the biofuel industry. In the first paper, we proposed a two-stage stochastic programming framework for the biofuel supply chain optimization problem considering uncertainties, including biomass supply availability, technology advancement, and biofuel market price. The results show that the stochastic factors have significant impacts on the decision on fast pyrolysis plant locations, especially when there is insufficient biomass. Also, farmers' participation can have a significant impact on the profitability and robustness of this supply chain design.

Another major challenge faced by the cellulosic biofuel industry is that investors are hesitant to take the risk to construct commercial scale production facilities. Techno-

economic analysis (TEA) has been widely adopted to overcome this challenge. The optimal facility locations and capacities as well as the logistic flow decisions for biomass supply and biofuel distribution should be incorporated into techno-economic analysis as well. In the second paper, the author aims to provide a new method that integrated the supply chain design into the techno-economic analysis as well by evaluating the economic feasibility of an integrated pathway on biomass pyrolysis and bio-oil gasification. The results indicate that hybrid fast pyrolysis and bio-oil gasification pathway is more suitable for a decentralized supply chain structure while biomass gasification pathway is more suitable for a single centralized facility supply chain structure.

Feeding millions of people throughout the world who face hunger every day is a formidable challenge. Precision agriculture has attracted increasing attention in the community of farmland management. Farmland management involves a sequence of planning and decision-making processes, including seed selection and irrigation schedule. In the third paper, a mixed integer programming optimization model is proposed to provide decision support on seed selection and irrigation water allocation for customized precision farmland management. The results show that significant increase of farmers' annual profit can be achieved by carefully choosing irrigation schedule and type of seed. The proposed model can also serve as a risk analysis tool for farmers facing seasonal irrigation water limits as well as a quantitative tool to explore the impact of precision agriculture.

The effect of limited water on corn grain yield is significant and management decisions are essential to optimize farmers' profits, particularly under stochastic environment. The fourth paper takes uncertainties such as crop price, irrigation water availability and precipitation amount into consideration. A multi-stage stochastic programming is formulated to evaluate the effects of structure of decision making process on farmers' income. The case study results indicate multi-stage stochastic programming is a promising way for farmland management under uncertainties and can increase farmers' income significantly.

In order to enhance the data utilization and results interpretation, statistical methods such as Monte-Carlo simulation considering parameter interactions, linear regression analysis, and moment matching method for scenario generation are also applied. The overarching goals of this dissertation is to quantify and manage the uncertainties along the modeling process and provide proper mechanisms that lead to optimal decisions. The outcomes of the research have the potential to accelerate the commercialization of second generation of biofuel and lead to sustainable utilization of water resources. The insights derived from the research contributed to the decision making process under uncertainties.

CHAPTER 1 GENERAL INTRODUCTION

1.1 Background

As a potential substitute for petroleum-based fuel, biofuels are playing an increasingly important role due to their economic, environmental, and social benefits. Biofuels include first generation biofuels made from sugar, starch, vegetable oil, etc., second generation biofuels made from non-food crops such as corn stover, switchgrass, forest, etc., and third generation biofuels mainly from algae. However, the 2007-2008 global food crisis was claimed to be related to biofuels production, and this food vs. fuel debate sets barriers for first generation biofuels from consumable grain and lipids [1]. Alternatively, the feedstocks for second generation biofuels are less land and water intensive, which will not result in significant negative impact on the food market [2]. On the other hand, farmland management under climate change and population growth is a pressing challenge that has become increasingly important due to food security considerations. Precision agriculture has attracted increasing attention in the community of farmland management. Over the years, the precision agriculture philosophy has enriched from simply "farming by soil" to a comprehensive system including irrigation planning, phenotypic selection, farm equipment guidance systems, product quality and environmental management etc. [3-5]. As the demand for agricultural products increases, water and arable land has become significant factors when considering agricultural production decisions. In summary, renewable energy and precision agriculture are emerging fields with increasing importance due to food, energy and water consideration.

US Environmental Protection Agency (EPA) revised the Renewable Fuel Standard in 2007, which aims to accelerate the domestic biofuel production and consumption. The RFS2 mandates that by the year 2022, at least 36 billion gallons per year of renewable fuels will be produced and blended into the transportation fuel, of which at least 16 billion gallons per year should be produced from cellulosic biomass feedstock [6]. However, the targeted cellulosic biofuel volume requirement for 2013 was revised down to be only 14 million gallons, which is significantly lower than the original target. This is mainly due to the high capital investment and logistic challenges in cellulosic biofuel. The supply chain activities of harvest, collection, storage, preprocessing, handling, and transportation represent one of the biggest challenges to the cellulosic biofuels industry, especially under siginificant uncertainties. Thus, it is timely and meaningful to study the economic feasibility of the commercialization of cellulosic biofuel considering the supply chain design under uncertainties.

Feeding millions of people throughout the world who face hunger every day is a formidable challenge. Precision agriculture has attracted increasing attention in the community of farmland management. Each year, farmers have to make decisions about what crops to plant. Farmers need to select the types of seeds and plan for irrigation carefully to achieve maximum profits. Thus, crop planning and irrigation water management on a farm scale are imperative for improved agricultural productivity and sustainable development [7]. The lack of decision making tools under uncertainties for renewable energy supply chain design and precision agriculture management serves the major motivation for this dissertation study.

One major challenge faced by the cellulosic biofuel industry is that investors are hesitatant to take the risk to construct commercial scale production facilities, and lack of facility cost information for the real production systems prohibit the improvement of production system to reduce costs and uncertainty [8]. Techno- economic analysis (TEA) has been widely adopted to overcome this challengue. There is an increasing literature on TEA for biofuels production pathways with a range of feedstock and final products [9-11]. However, the process design and techno-economic analysis of the integrated pathway have not been studied extensively.

There has been a growing body of literature on crop rotations at a regional scale [12, 13], land use patterns, and policy and environment issues on a farm scale [14]. Mathematical programming has been widely used in farmland management and supply chain network design. Shah [15] reviewed the previous studies in modeling, planning, and scheduling with some real world examples to summarize the challenges and advantages of supply chain optimization. Eksioglu et al [16] formulated a model to determine the numbers, locations, and capacities of the biorefineries, and conducted a case study for Mississippi in the U.S. to illustrate and verify the optimization model. Sethi et al [7] modeled a linear programming problem to find maximum annual net return under different soil types, cropping patterns, and types of agriculture. One of the biggest challenges of food and energy industry is the decision making under uncertainties. Most of the literature assumes all the parameters in the system are deterministic. However, these industries are highly affected by the uncertainties such as market price, biomass yield, farmers' participation, and technology advancement. As a result, it is of vital importance to consider the uncertainties in the decision making process.

1.2 Introduction of Individual Components

The individual components of this dissertation are introduced in more detail in this section. In the first paper, an advanced biofuels supply chain is proposed to reduce biomass transportation costs and take advantage of the economics of scale for a gasification facility. In this supply chain, biomass is converted to bio-oil at widely distributed small-scale fast pyrolysis plants, and after bio-oil gasification, the syngas is upgraded to transportation fuels at a centralized biorefinery. A two-stage stochastic programming is formulated to maximize biofuel producers' annual profit considering uncertainties in the supply chain for this pathway. The first stage makes the capital investment decisions including the locations and capacities of the decentralized fast pyrolysis plants as well as the centralized biorefinery, while the second stage determines the biomass and biofuels flows.

A case study of Iowa is presented to illustrate and validate this supply chain design and optimization model. The results show that uncertain factors such as biomass availability, technology advancement, and biofuel price can be pivotal in this supply chain design and optimization. The locations of fast pyrolysis plants and logistic decisions are sensitive to uncertainties while the capacity levels are insensitive. The stochastic model outperforms the deterministic model in the stochastic environment, especially when there is insufficient biomass. In addition, farmers' participation has a significant impact on the decision making process. It is appropriate and necessary to apply a stochastic programming framework to deal with the uncertainties, especially at a low farmers' participation level. As farmers' participation increases, the supply chain design and optimization model will become more profitable and more robust against the uncertainties along the supply chain.

In the second paper, a techno-economic analysis method considering logistic configurations is proposed. The economic feasibility of a low temperature biomass gasification pathway and an integrated pathway with fast pyrolysis and bio-oil gasification are evaluated and compared with the proposed method in Iowa. The results show that both pathways are profitable, biomass gasification pathway could achieve an Internal Rate of Return (IRR) of 10.00% by building a single biorefinery and integrated bio-oil gasification pathway could achieve an IRR of 3.32% by applying decentralized supply chain structure. The supply chain analysis results show BMG pathway is more economically feasible than BOG pathway in Iowa when realistic supply chain configurations and constraints are considered. Different production pathways could have its preferred supply chain structure. BOG pathway is more suitable for a decentralized supply chain structure while BMG pathway is more suitable for a single facility supply chain structure. The supply chain configuration demonstrates the trade-off between feedstock shipping cost and the capital investment of multiple facilities in different scenarios.

The sensitivity analysis shows that the MSP is most sensitive to internal rate of return, fuel yield, biomass feedstock cost, and fixed capital investment. A Monte-Carlo simula- tion considering interactions among parameters is also proposed and conducted. Both cases in the Monte-Carlo simulation results for single 2000 MT/D facility show that the range of MFSP is about 4–7 $/ GGE for BMG pathway. These results indicate that even through BMG pathway has better economic performance than BOG path- way, both pathways are at high risk at this point. In addition, assumptions for distribution as well as its variance covariance structure can take significant impact on the uncertainty analysis.

In the third paper, a farm-level precision farmland management model based on mixed integer linear programming is proposed. Farmland management involves several planning and decision making tasks including seed selection and irrigation management. Optimal decisions are designed for pre-season planning of crops and irrigation water allocation. The model captures the effect of size and shape of decision scale as well as special irrigation patterns. The authors illustrate the model by a case study based on a farm in California, the U.S. and show the model is economically optimal and flexible. The results show that threefold increase of annual net profit for farmers could be achieved by carefully choosing irrigation and seed selection. Although farmers could increase profits by applying precision management to seed or irrigation alone, profit increase is more significant if farmers apply precision management on seed and irrigation simultaneously.

The proposed model can also serve as a risk analysis tool for farmers facing seasonal irrigation water limits as well as a quantitative tool to explore the impact of precision agriculture.

The fourth paper is an extention of the third paper by considering uncertainties such as crop price, irrigation water availability, and precipitation amount. A multi-stage stochastic programming is formulated to maximize farmer's annual profit. The first stage decisions including the seed type selection and plant population selection, while the later stage determine the irrigation schedule. The case study based on a farm in Nebraska show that taking corn price, precipitation amount, and irrigation water availability uncertainties into consideration can increase farmer’s profit. In the stochastic programming results, more conservative first stage decisions are made such as select high drought resistance seed. These decisions preform more robust in the stochastic environment. These results indicate multi-stage stochastic programming is a promising way for farmland management under uncertainties and can increase farmers’ income significantly.

1.3 Dissertation Structure

The remainder of the dissertation is organized as follows. The first paper on supply chain design under uncertainty for advanced biofuel production based on bio-oil gasification is present in Chapter 2 and has been published in Energy [17]. In Chapter 3, we present the second paper on techno-economic analysis of biofuel production considering logistic configurations and has been published in Bioresource Technology [18]. In Chapter 4, we present the third paper on a Farm-level Precision Land Management

Framework Based on Integer Programming. This paper has been published in PLOS ONE. In Chapter 5, we propose a multi-stage stochastic programming for farmland management under uncertainties. This chapter of dissertation is preparing to submit to European Journal of Operational Research. Finally, Chapter 6 concludes the dissertation and proposed possible future research directions.

CHAPTER 2 SUPPLY CHAIN DESIGN UNDER UNVERTAINTY FOR ADVANCED BIOFUEL PRODUCTION BASED ON BIO-OIL GASIFICATION[1]

Abstract

An advanced biofuels supply chain is proposed to reduce biomass transportation costs and take advantage of the economics of scale for a gasification facility. In this supply chain, biomass is converted to bio-oil at widely distributed small-scale fast pyrolysis plants, and after bio-oil gasification, the syngas is upgraded to transportation fuels at a centralized biorefinery. A two-stage stochastic programming is formulated to maximize biofuel producers' annual profit considering uncertainties in the supply chain for this pathway. The first stage makes the capital investment decisions including the locations and capacities of the decentralized fast pyrolysis plants as well as the centralized biorefinery, while the second stage determines the biomass and biofuels flows. A case study based on Iowa in the U.S. illustrates that it is economically feasible to meet desired demand using corn stover as the biomass feedstock. The results show that the locations of fast pyrolysis plants are sensitive to uncertainties while the capacity levels are insensitive. The stochastic model outperforms the deterministic model in the stochastic environment, especially when there

[1] This chapter of dissertation has been published in Energy

is insufficient biomass. Also, farmers' participation can have a significant impact on the profitability and robustness of this supply chain.

2.1 Introduction

As a potential substitute for petroleum-based fuel, biofuels are playing an increasingly important role due to their economic, environmental, and social benefits. However, the 2007-2008 global food crisis was claimed to be related to biofuels production [1], and this food vs. fuel debate set barriers for first generation biofuels from consumable grain and lipids. Alternatively, second generation biofuels are produced from nonedible plant residues or dedicated energy crop, such as corn cobs, corn stover, switchgrass, miscanthus, and woody biomass. As a result, the feedstocks for second generation biofuels are less land and water intensive, which will not result in significant negative impact on the food market [2]. According to the revised Renewable Fuel Standard (RFS2) established in 2007, at least 36 billion gallons per year of renewable fuels will be produced by 2022 in the U.S., of which at least 16 billion gallons per year will be from cellulosic biofuels [19]. However, the targeted cellulosic biofuel volume requirement for 2013 was revised to be only 14 million gallons, which is significantly lower than the original target. This is mainly due to the high capital investment and logistic challenges in cellulosic biofuel. The supply chain activities of harvest, collection, storage, preprocessing, handling, and transportation dealing with uncertainties represent one of the biggest challenges to the cellulosic biofuels industry. Thus, it is timely and meaningful to study the economic feasibility of the commercialization of cellulosic biofuel considering the supply chain design under uncertainties.

Biomass can be converted to transportation fuels through a variety of production pathways, including biochemical and thermochemical platforms. One example of biochemical pathways is corn ethanol production from fermentation. Another example is the thermochemical conversion of biomass to produce transportation fuels, which has recently moved to the forefront of biofuel research and development. Fast pyrolysis and gasification are two of the most prominent technologies for thermochemical conversion of cellulosic biomass.

Fast pyrolysis thermally decomposes organic compounds in the absence of oxygen, and the products include bio-oil, bio-char, and non-condensable gases [20]. Fast pyrolysis reactors typically run at temperatures between 400 ºC and 600 ºC and can produce approximately 70% (by weight) bio-oil [21]. The other 30% is split between non-condensable gases (e.g., carbon dioxide or methane) and bio-char. The non-condensable gases and bio-char could be combusted to provide heat for the facility. In addition, bio-char is mostly organic carbon which can be sequestered or gasified to produce syngas [22]. Bio-oil has three to five times the energy density of raw biomass [23]. However, due to the high viscosity and acidity, bio-oil needs to be upgraded to be used as transportation fuels. The bio-oil upgrading has proven to be a challenging process due to low conversion efficiency and fuel quality. Unlike fast pyrolysis, biomass gasification runs at a much higher temperature (800 ºC - 1300 ºC) and is a relatively mature technology. The syngas produced from the biomass gasification process will typically go through the Fischer-Tropsch synthesis to produce liquid transportation fuels [1]. However, commercialization of biomass gasification has been hampered by its high capital and operating costs due to the challenges of transporting bulky solid biomass over a long distance, processing solid feedstock at high pressure, and removing contaminants from the product gas stream. The techno-economic analysis of biomass gasification by Swanson et al. claims that the minimum fuel selling price is $4-5 per gallon of gasoline equivalent and the capital investment requirement is $500-650 million for a 2000 metric tons per day facility [9].

It is thus necessary to reduce system cost and improve supply chain efficiency to improve the economic feasibility and competitiveness of the advanced biofuel production pathways. To reduce feedstock transportation cost, it has been suggested that biomass can be converted to bio-oil via fast pyrolysis near harvest sites, and then the bio-oil can be transported to an upgrading plant for transportation fuels production [24]. In this paper, the proposed hybrid production pathway is to combine the two prominent thermochemical production pathways. Biomass fast pyrolysis produces bio-oil in relatively small processing plants at distributed locations so that the transportation of bulky biomass over a long distance can be avoided. After mild hydrotreating, the bio-oil is then transported to a centralized gasification facility to produce transportation fuels. This pathway could also simplify syngas cleanup as ashes in biomass played a significant role in the gasification process [25]. It should be recognized that a centralized plant has advantages such as economies of scale, an inventory buffer storage reduction, and administration overhead cost savings [26].

One of the biggest challenges of the advanced biofuel production industry is the design of supply chain networks under uncertainties. There is rich literature on supply chain network design. Shah reviewed the previous studies in modeling, planning, and scheduling with some real world examples to summarize the challenges and advantages of supply chain optimization [15]. An et al. compared the supply chain research of petroleum-based fuel and biofuel [27]. Eksioglu et al. formulated a model to determine the numbers, locations, and capacities of the biorefineries, and conducted a case study for Mississippi in the U.S. to illustrate and verify the optimization model [16]. Nixon et al. used a goal programming model to deploy a pyrolysis plants supply chain in Punjab, India [28]. Most of the literature on biofuel supply chain design assumes all the parameters in the system are deterministic. However, the biofuel industry is highly affected by the uncertainties along the supply chain such as biomass supply availability, technology advancement, and biofuel price. For example, the biomass feedstock supply is highly dependent on biomass yield and farmers' participation. As a result, it is of vital importance to design the biofuel supply chain considering the uncertainties along the supply chain. Kim et al. considered a two-stage stochastic model using bounds of the parameters to determine the capacities and locations of the biorefineries [29]. Alex et al. formulated a mixed integer linear programming model to determine optimal locations and capacities of biorefineries [30]. Osmani et al. used stochastic optimization to deal with the uncertainties in biomass yield and price as well as biofuel demand and price [31]. Since thermochemical pathways to produce cellulosic biofuel is a relatively recent technology advancement, decentralized supply chain design have not been studied extensively, especially scenario under uncertainties. This paper aims to provide a mathematical programming framework with a two-stage stochastic programming approach to design the supply chain network considering uncertainties along the supply chain. The production pathway under consideration is bio-oil gasification, with bio-oil production from biomass fast pyrolysis at decentralized facilities and syngas production and fuel synthesis in a centralized gasification facility. This model provides methodological insights for decision makers on the capital investment decisions and logistic decisions for the biofuel supply chain.

The remainder of the paper is organized as follows: in Section 2.2, the problem statement for the biofuel supply chain design is presented. Then, we discuss the deterministic mixed integer linear programming model and the two-stage stochastic programming models in Section 2.3. A case study of Iowa is conducted to illustrate and validate the optimization model in Section 2.4. Finally, we conclude the paper in Section 2.5 with summary and potential research directions.

2.2 Problem Statement

As mentioned, one of the most important decisions faced by the biofuel industry is the design of the supply chain networks, especially under system uncertainties. This provides the major motivation for this study.

The supply chain system schematics for the bio-oil gasification pathway are shown in Figure 2.1. Biomass is collected and consolidated at the county level. Biomass is collected and consolidated at the county level. Biomass is then transported to the decentralized fast pyrolysis facilities to be converted to bio-oil. Mild-hydrotreated bio-oil is transported to a centralized gasification facility to produce transportation fuels. It is assumed that each biomass feedstock supply location/county can serve multiple fast pyrolysis facilities, and that each fast pyrolysis facility can acquire feedstock from multiple biomass supply locations. The locations for the decentralized fast pyrolysis facilities and centralized gasification facility are assumed to be the centroids of counties.

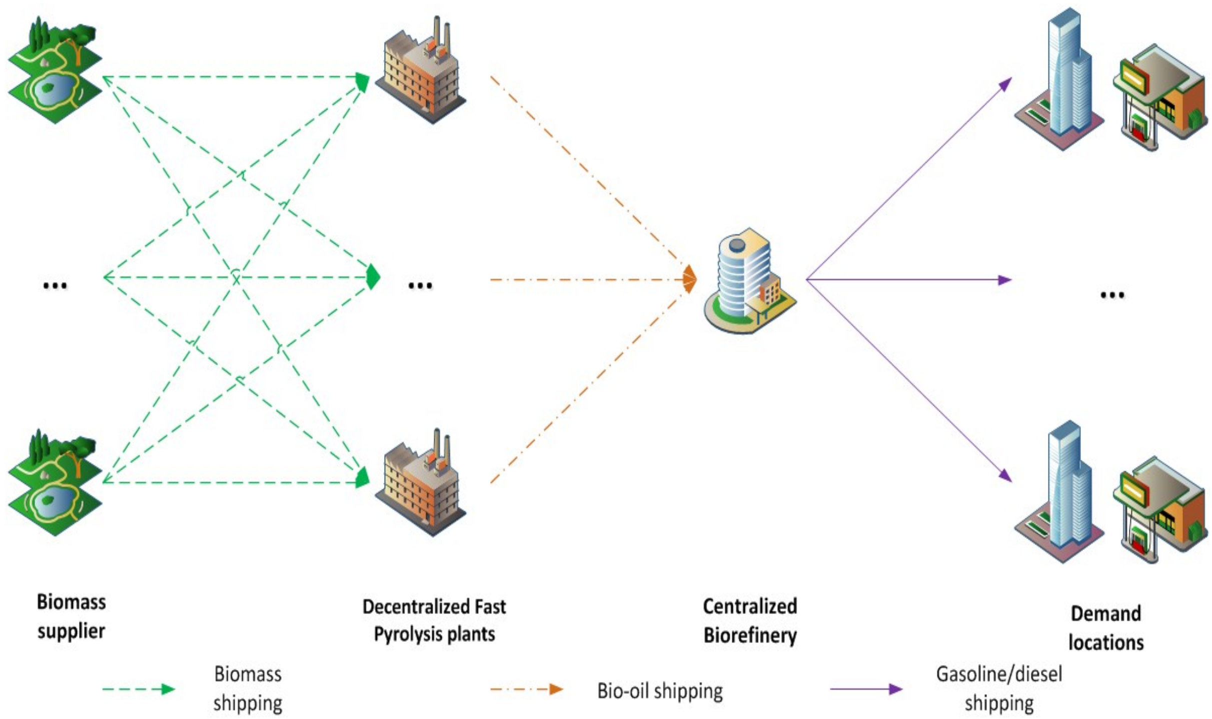


Figure 2.1 System schematics of supply chain

The supply chain network design of biofuel production is highly affected by uncertainties along the supply chain such as biomass supply availability, technology advancement, and biofuel price. The biomass supply availability is highly dependent on crop yields and farmers' participation, the conversion rates are affected by technology advancement and operating conditions, and the biofuel price would change based on market conditions and enacted policies. Thus, it is of vital importance to make the supply network design decisions with system uncertainties taken into consideration. Stochastic programming is one of the most widely used modeling frameworks to study decision making under uncertainties.

The goal of this paper is to provide a two-stage stochastic programming framework for the biofuel supply chain optimization problem considering uncertainties. The comparison and analysis of the results provide methodological suggestions on capital investment and logistic decisions. The insights derived from this study can contribute to the body of knowledge in decision making under uncertainties.

2.3 Model Formulation

In this section, we introduce the deterministic and stochastic models for this biofuel supply chain design problem. The objective is to maximize the annual profit in a biofuel network based on the hybrid production pathway of bio-oil gasification. The deterministic mixed integer linear programming model is firstly introduced as a baseline model and then the two-stage stochastic model is presented to address the uncertainties in the supply chain design problem. The stochastic programming framework bears the concept of recourse, which means some decisions (recourse actions) are taken after uncertainties have been realized. In other words, first-stage decisions are made by taking some factors' future effects into account. In the second stage, the actual value of the variables becomes known and some corrective actions can be taken [32].

2.3.1 Mathematical notations

The mathematical notations are summarized in Table 2.1.

Table 2.1 Notations for deterministic model

Subscripts		
$\boldsymbol{i}$	$1,2,\dots,I$	Biomass supply locations
$\boldsymbol{j}$	$1,2,\dots,J$	Candidate fast pyrolysis facility locations
$\boldsymbol{k}$	$1,2,\dots,K$	Biofuel demand locations
$\boldsymbol{l}$	$1,2,\dots,L$	Allowed fast pyrolysis capacity levels
$\boldsymbol{m}$	$1,2,\dots,M$	Candidate refining facility locations
Decision Variables		
$\boldsymbol{x_{ij}}$	Amount of biomass transported from supply location i to candidate fast pyrolysis facility location j	
$\boldsymbol{y_{jm}}$	Amount of bio-oil transported from candidate fast pyrolysis facility location j to candidate refining facility location m	
$\boldsymbol{z_{mk}}$	Amount of biofuels transported from refining facility location m to demand location k	
$\boldsymbol{a_{jl}}$	Whether a fast pyrolysis facility of capacity level l is planned at candidate facility location j (binary variable)	
$\boldsymbol{g_m}$	Whether a refining facility is planned at candidate refining facility location m (binary variable)	
Parameters		
B	Total budget	
$\boldsymbol{C^{UP}}$	Capital cost of the centralized refining facility	
$\boldsymbol{C_l^{Cap}}$	Capital cost of the decentralized fast pyrolysis facility at level l	
$\boldsymbol{P_k}$	Biofuels price at demand location k	
$\boldsymbol{D_k}$	Biofuels demand at demand location k	
$\boldsymbol{Pe_k}$	Penalty for not meeting the demand at demand location k	
$\boldsymbol{Pe_k}'$	Penalty for exceeding the demand at demand location k	
$\boldsymbol{C_i^{Col}}$	Unit biomass collecting cost at supply location i	
$\boldsymbol{C^{MO}}$	Unit conversion cost from dry biomass to bio-oil	
$\boldsymbol{C^{OF}}$	Unit conversion cost from bio-oil to biofuels	
$\boldsymbol{C_{ij}^{BM}}$	Unit biomass shipping cost from supply location i to candidate fast pyrolysis facility location j	
$\boldsymbol{C_{jm}^{BO}}$	Unit bio-oil shipping cost from candidate fast pyrolysis facility location j to candidate refining facility location m	
$\boldsymbol{C_{mk}^{BF}}$	Unit biofuel shipping cost from candidate refining facility location m to demand location k	

Table 2.1 continued

$\boldsymbol{U_l}$	Capacity of fast pyrolysis facility at level l
$\boldsymbol{V}$	Capacity of refining facility
$\boldsymbol{S_i}$	Available biomass feedstock at location i
$\boldsymbol{\alpha}$	Sustainability factor
$\boldsymbol{\beta}$	Conversion factor from wet biomass to dry biomass
$\boldsymbol{\gamma}$	The loss factor of biomass during collection and transportation
$\boldsymbol{\theta_1}$	Conversion ratio, metric ton of bio-oil per metric ton of dry biomass
$\boldsymbol{\theta_2}$	Conversion ratio, metric ton of biofuels per metric ton of bio-oil
$\boldsymbol{\delta}$	Availability factor

2.3.2 Deterministic model

In the deterministic mixed integer linear programming model, all the system parameters are assumed to be known with certainty.

The objective function is to maximize the annual profit, which can be defined as the revenue from selling the biofuels subtracted by the total system costs along the supply chain including the potential penalties. Penalties are imposed on the unmet demand which is based on the assumption that the producers have to purchase fuels form other sources to satisfy unmet demand. Penalties are also imposed for the surplus production due to additional inventory holding and storage costs. A variety of system costs have been considered in the model including facility capital investment cost, biomass collection cost, biofuel conversion cost, and logistics cost.

Firstly, the total capital cost for the decentralized fast pyrolysis facility at level l is $\sum_{j=1}^{J}\sum_{l=1}^{L} C_l^{Cap} a_{jl}$. With the assumption that the facilities have an n-year operation life and an interest rate of i, the annual amortized capital cost is $\frac{i(i+1)^n}{(i+1)^n-1}(\sum_{j=1}^{J}\sum_{l=1}^{L} C_l^{Cap} a_{jl} + C^{UP})$. Secondly, the cost of collection biomass from different feedstock location is $\sum_{i=1}^{I}\sum_{j=1}^{J} C_i^{Col} x_{ij}$. Thirdly, $C^{MO}(1-\gamma)\beta\sum_{i=1}^{I} x_{ij}$ is the fast pyrolysis conversion cost from biomass to bio-oil and $C^{OF}\sum_{j=1}^{J}\sum_{m=1}^{M} y_{jm}$ is the conversion cost from bio-oil to biofuel at the gasification and upgrading biorefinery. Lastly, the logistics costs include the biomass shipping cost from biomass feedstock locations to fast pyrolysis facility locations, the bio-oil shipping cost from fast pyrolysis facility locations to gasification and upgrading biorefinery location, and the biofuel shipping cost from gasification and upgrading biorefinery location to demand locations.

In sum, the objective function can be formulated as follows:

$$
\begin{aligned}
max\, \zeta &= income - penalty - cost \\
&= \sum_{k=1}^{K}(P_k \sum_{m=1}^{M} z_{mk}) - \{\left(D_k - \sum_{m=1}^{M} z_{mk}\right)_+ * Pe_k + \left(\sum_{m=1}^{M} z_{mk} - D_k\right)_+ \\
&* Pe_k'\} - \{\frac{i(i+1)^n}{(i+1)^n - 1}(\sum_{j=1}^{J}\sum_{l=1}^{L} C_l^{Cap} a_{jl} + C^{UP}) + \sum_{i=1}^{I}\sum_{j=1}^{J} C_i^{Col} x_{ij} \\
&+ C^{MO}(1-\gamma)\beta \sum_{i=1}^{I} x_{ij} + C^{OF}\sum_{j=1}^{J}\sum_{m=1}^{M} y_{jm} + \sum_{i=1}^{I}\sum_{j=1}^{J} C_{ij}^{BM} x_{ij} \\
&+ \sum_{j=1}^{J}\sum_{m=1}^{M} C_{jm}^{BO} y_{jm} + \sum_{m=1}^{M}\sum_{k=1}^{K} C_{mk}^{BF} z_{mk}\}
\end{aligned}
$$

The constraint (1) is included to ensure that the sum of capital cost of decentralized fast pyrolysis facilities and centralized biorefinery does not exceed the total budget.

$$B \geq C^{UP} + \sum_{j=1}^{J}\sum_{l=1}^{L} C_l^{Cap} a_{jl} \tag{1}$$

The total amount of biomass transported from supply location i to all the candidate fast pyrolysis facility locations should not exceed the available feedstock at that supply location as denoted in constraint (2). α is the sustainability factor which is the percentage of biomass that has to leave in the field to sustain the soil nutrients. δ is the availability factor which is defined as the ratio of the available biomass to collectable biomass. This factor represents the social factors that could impact the biomass availability for biofuel production such as farmers' willingness to participate [33].

$$\sum_{j=1}^{J} x_{ij} \leq (1-\alpha)\delta S_i, \ \forall i \quad (2)$$

The facility capacity limits are included in the model in constraint (3) and constraint (4). The loss factor $\gamma \in [0,1)$ is the fraction weight loss of biomass during the collection, transportation, and unloading process and β is the conversion ratio from wet biomass to dry biomass on the weight basis.

$$\sum_{l=1}^{L} U_l a_{jl} \geq (1-\gamma)\beta \sum_{i=1}^{I} x_{ij}, \forall j \quad (3)$$

$$V g_m \geq \sum_{j=1}^{J} y_{jm}, \ \forall m \quad (4)$$

There should be no more than one fast pyrolysis facility planned in each candidate facility location as shown in constraint (5). In addition, only one centralized refining facility will be constructed in one region of interest (typically one state) as denoted in constraint (6).

$$\sum_{l=1}^{L} a_{jl} \leq 1, \forall j \quad (5)$$

$$\sum_{m=1}^{M} g_m = 1 \quad (6)$$

We assume that biomass is converted to bio-oil with conversion efficiency θ_1 and bio-oil is converted to biofuel with conversion efficiency θ_2 on the weight basis. Thus, we have the following conversion balance constraints (7) and (8):

$$(1-\gamma)\beta\theta_1 \sum_{i=1}^{I} x_{ij} = \sum_{m=1}^{M} y_{jm}, \forall j \tag{7}$$

$$\theta_2 \sum_{j=1}^{J} \sum_{m=1}^{M} y_{jm} = \sum_{m=1}^{M} \sum_{k=1}^{K} z_{mk} \tag{8}$$

In summary, this mixed integer linear programming model aims to maximize the annual profit considering the capital investments and logistics decisions. This deterministic model provides the baseline for the stochastic programming model in the next sections.

2.3.3 Two-stage stochastic programming model

Feedstock availability, fuel price, capital costs, logistic costs, and technology advancement are among the most influential stochastic parameters along the biofuel supply chain [34]. These uncertainties can be incorporated into the stochastic modeling framework to assist decision making.

In this section, the two-stage stochastic programming model is discussed considering the uncertainties of the biomass availability, technology advancement, and biofuel prices. The stochastic parameters in this model are assumed to be discretely distributed. We use subscript s to represent scenario with corresponding probability Pr_s

and this subscript is also incorporated into the decision variables and parameters. The two-stage stochastic programming model is formulated as follows:

$$
\begin{aligned}
max\,\zeta = -\frac{i(i+1)^n}{(i+1)^n-1}\sum_{j=1}^{J}\sum_{l=1}^{L} C_l^{Cap} a_{jl} + \sum_{s=1}^{S} Pr_s\{\sum_{k=1}^{K}\sum_{m=1}^{M}(P_{ks} z_{mks}) \\
-(\left(D_k - \sum_{m=1}^{M} z_{mks}\right)_+ \times Pe_k + \left(\sum_{m=1}^{M} z_{mks} - D_k\right)_+ \times Pe'_k) \\
-(\sum_{i=1}^{I}\sum_{j=1}^{J} C_i^{Col} x_{ijs} + C^{MO}(1-\gamma)\beta\sum_{i=1}^{I} x_{ijs} + C^{OF}\sum_{j=1}^{J}\sum_{m=1}^{M} y_{jms} \\
+(\sum_{i=1}^{I}\sum_{j=1}^{J} C_{ij}^{BM} x_{ijs} + \sum_{j=1}^{J}\sum_{m=1}^{M} C_{jm}^{BO} y_{jms} \sum_{m=1}^{M}\sum_{k=1}^{K} C_{mk}^{BF} z_{mks}))\}
\end{aligned}
$$

$s.t.$

$Constraints$ (1), (5), (6).

$\sum_{j=1}^{J} x_{ijs} \le (1-\alpha)\delta S_{is}, \forall i, \forall s$ (9)

$\sum_{l=1}^{L} U_l a_{jl} \ge (1-\gamma)\beta \sum_{i=1}^{I} x_{ijs}\,, \forall j, \forall s$ (10)

$V g_m \ge \sum_{j=1}^{J} y_{jms}\,, \forall m, \forall s$ (11)

$(1-\gamma)\beta\theta_{1,s}\sum_{i=1}^{I} x_{ijs} = \sum_{m=1}^{M} y_{jms}\,,\ \forall j, \forall s$ (12)

$\theta_{2,s}\sum_{j=1}^{J}\sum_{m=1}^{M} y_{jms} = \sum_{m=1}^{M} z_{mks}\,,\ \forall s$ (13)

$x_{ijs}, y_{jms}, z_{mks} \ge 0, a_{jl}, g_m \in \{0,1\}, \forall i, j, k, m, l, s$ (14)

The first-stage decisions involve variables which should be decided before the uncertainties are realized. After the uncertainties are realized, the second-stage decisions are made. In this supply chain network design model, the first-stage decision variables include the binary variables a_{jl} and g_m, which make the capital investment decisions including the facility locations (decentralized fast pyrolysis and centralized refining facilities) and capacities of the decentralized fast pyrolysis facilities. The second-stage decision variables x_{ijs}, y_{jms}, and z_{mks} determine the biomass and biofuels flows.

Constraints (1), (5), and (6) are the first-stage constraints, these constraints remain the same in all scenarios and they are same as in the deterministic linear program model. The rest of the constraints change based on the stochastic scenario. The rest of the constraints change based on the stochastic scenario. Note that this model is a generic method to deal with uncertainties in a supply chain and can be adapted to other types of uncertainties and supply chain settings.

One of the most commonly used methods for scenario generation is the moment matching method. This method aims to construct a set of scenarios with corresponding probabilities such that the statistical properties of the approximating distribution match the specified statistical properties based on historical data or reality. This is achieved by minimizing the differences between the statistical properties of the constructed distribution and the known specifications, subject to nonnegative probabilities that sum up to one [35].

2.4 Case Study

We apply the supply chain design framework for a case study based on Iowa in the U.S. to illustrate and validate the optimization model. Iowa possesses the largest quantity of corn stover in the United States and has been one of the leading states of corn ethanol and soybean biodiesel production [36]. With an abundance of cellulosic biomass, Iowa has potential for cellulosic biofuel production via thermochemical conversion processes.

2.4.1 Data sources

The centroids of 99 counties of Iowa are chosen as candidate biomass (corn stover in this case study) supply locations, the potential sites for distributed fast pyrolysis facilities, and the candidate location for the centralized gasification facility. The annual corn stover yield is estimated based on corn grain yield with the residue harvest index of 0.5 (i.e., 50% of the dry mass of the corn plant is grain and the rest 50% is stove) [37]. The weight of #2 corn at 15.5% moisture is applied to calculate the corn grain yields [38]. The county level corn production and yield data from 2003-2012 are collected from the National Agricultural Statistics Service (NASS), United States Department of Agriculture

(USDA) [39]. The average county level corn stover yield in Iowa for 2003-2012 is shown in Figure 2.2 with the darkness of the shade corresponding to the corn stove yield.

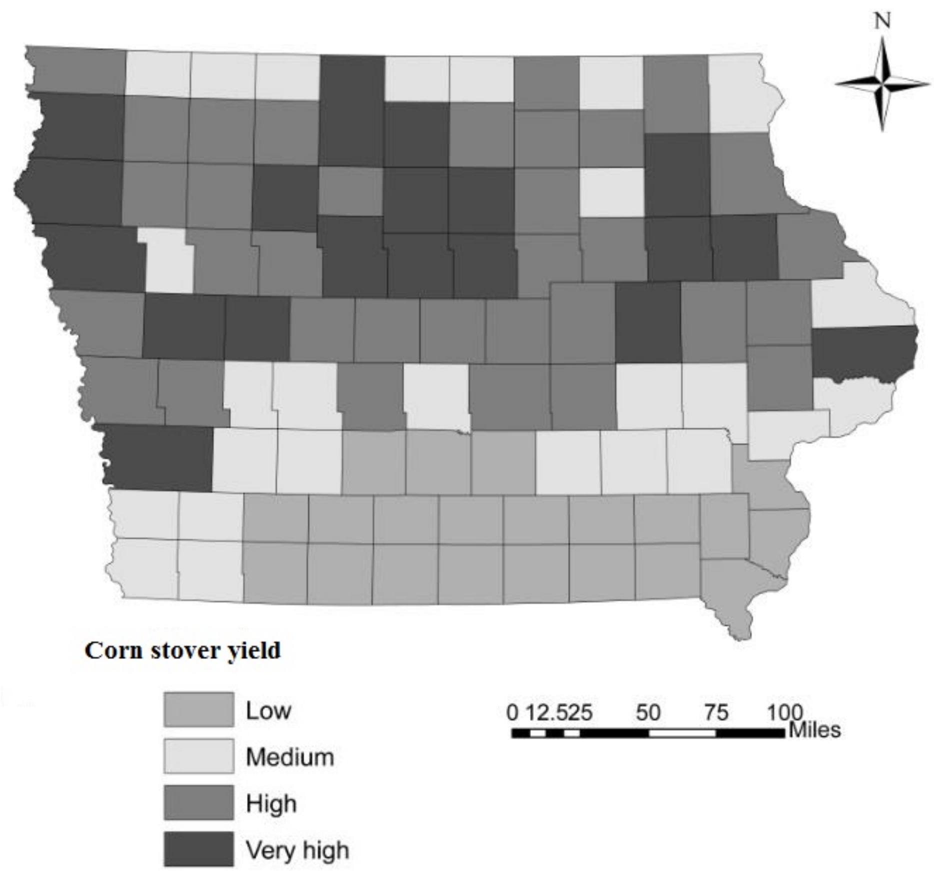


Figure 2.2 Average corn stover yield in Iowa (2003-2012)

In addition, the collectable corn stover is limited by growing conditions, soil nutrient levels, and method of harvest. Montross et al. reported the collection efficiencies of using three strategies in Kentucky: bale only to be 38%; rake and bale to be 55%; and mow, rake, and bale to be 64% [40]. Schechinger and Hettenhaus reported collection efficiencies of 40% to 50% without raking and 70% with raking in large-scale stover collection operations in Nebraska and Wisconsin [41]. Lindstrom suggested that a 30% removal rate would not significantly increase soil loss [42]. Later, Papendick et al. shows that a 30% removal rate results in 93% soil cover after residue harvest [43]. The National Resource Conservation Service (NRCS) suggests that a minimum of 30% of stover cover must remain in the field to prevent soil erosion [44]. In this analysis, we assume the sustainability factor to be 0.3, which means at least 30% of the stover must be left in the field to promote soil health.

The collection cost for corn stover is different for each county due to the differences in collection quantities and collection methods. The collection cost utilized in this case study is based on the regression analysis from Graham et al. [45]. Biomass loss factor, which accounts for possible mass loss during loading, transportation, and unloading of the biomass, is assumed to be 0.05 in this analysis.

The total gasoline demand of Iowa is based on the state-level gasoline consumption data from the Energy Information Administration (EIA) [46]. Weekly retail gasoline prices for the Midwest area from 2003 to 2012 are also from EIA [47]. Gasoline demand of each demand area is assumed to be proportional to the population of metropolitan statistical areas (MSAs). The partitions and population information of Iowa MSAs are based on U.S. Census Bureau [48].

All the biomass suppliers, biorefineries, and demand locations are assumed to be at the county centroids. Transportation distances for biomass, bio-oil and biofuels are calculated using the great circle distance, which is defined as the shortest distance between the two locations on a sphere surface. In addition, the actual distances have been adjusted to account for the difference in the transportation methods by the circuit factors from the Congressional Budget Office [49].

The fixed transportation cost of corn stover via truck is $5.34/metric ton and the variable cost of $0.23/metric ton-mile [50]. The transportation cost of bio-oil via truck is assumed to be equal to the national average truck shipping cost of $0.312/metric ton-mile based on Bureau of Transportation Statistics (BTS). The transportation cost of biofuel via pipeline is assumed to be equal to the national average oil pipeline cost, which is $0.032/metric ton-mile [51]. The cost data have been adjusted to the 2012 US dollars.

In the fast pyrolysis process, the biomass is converted into bio-oil (53-78%), char (12-34%), and gas (8-20%) [52]. The bio-oil yield is assumed to follow the normal distribution based on the experimental results from Iowa State University. In this study, the fluidized bed reactor is employed in the fast pyrolysis which has an average conversion ratio of 0.63 from biomass to bio-oil on weight basis [53]. The conversion ratio from bio-oil to biofuel is not available due to lack of experimental data. Limited experiment shows high carbon conversion of gasification but low efficiency from syngas to fuel (due to the diverse H_2/CO ratio). Raffelt et al. reported a conversion ratio of 0.156 on weight basis for slurry (80% bio-oil and 20% char) gasification [52]. We assume that the conversion ratio from bio-oil to biofuel follows a normal distribution with an average of 0.20 on weight basis. With these assumptions, the average fuel yield for the pathway under analysis would be 31.2 million GGE per year for the plant size to of 2000 metric ton biomass per day facility. This is consistent with reported fuel yield of 29.3-58.2 million GGE per year for 2000 metric ton per day facility [54].

Wright et al. reported that the capital cost of centralized gasification plant with a capacity of 550 million GGE per year is about 1.47 billion [55]. The capital cost of distributed fast pyrolysis facility with a capacity of 2,000 metric ton per day is $200 million [53]. The commonly used scaling factor of 0.6 (the "sixth-tenth rule") is applied to estimate capital cost for facilities with other capacity levels. In this study, we consider three capacity levels of distributed fast pyrolysis facilities: 500, 1000, and 2000 metric ton per day. According to RFS2, at least 36 billion gallons per year of renewable fuels will be produced by 2022, which is about 28% of the national gasoline consumption. In this study, we assume the centralized gasification and upgrading plant has a capacity of 550 million GGE per year, which could satisfy more than 30% of the gasoline consumption in Iowa. Thus, we only need to consider one centralized bio-oil gasification and upgrading facility in this case study.

It is assumed that all the facilities have a 20-year operation life and an interest rate of 10%; the online time of all the facilities is 328 days per year (equivalent capacity factor of 90%). In the following two sections, the computational results of the biofuel supply chain design for both deterministic case and stochastic case are presented.

2.4.2 Analysis for deterministic case

In the deterministic case, 17 distributed fast pyrolysis plants will be built, and all of them are at the highest capacity level (2000 metric ton per day). This is mainly due to the budget limit and economies of scale. The centralized gasification plant is planned to be located in Hamilton County. The optimal locations for these facilities are shown in

Figure 2.3.The shaded areas are biomass feedstock suppliers (71 counties) in this case. These counties are mainly located at the central and northern part of Iowa, which have a higher yield of corn and thus have better availability for corn stover. Several previous studies [17, 56] showed similar site selection decisions, but there are more biomass feedstock counties involved in our case. The counties' locations of distributed fast pyrolysis plants illustrate the trade-off between biomass collection as well as transportation cost and bio-oil transportation cost.

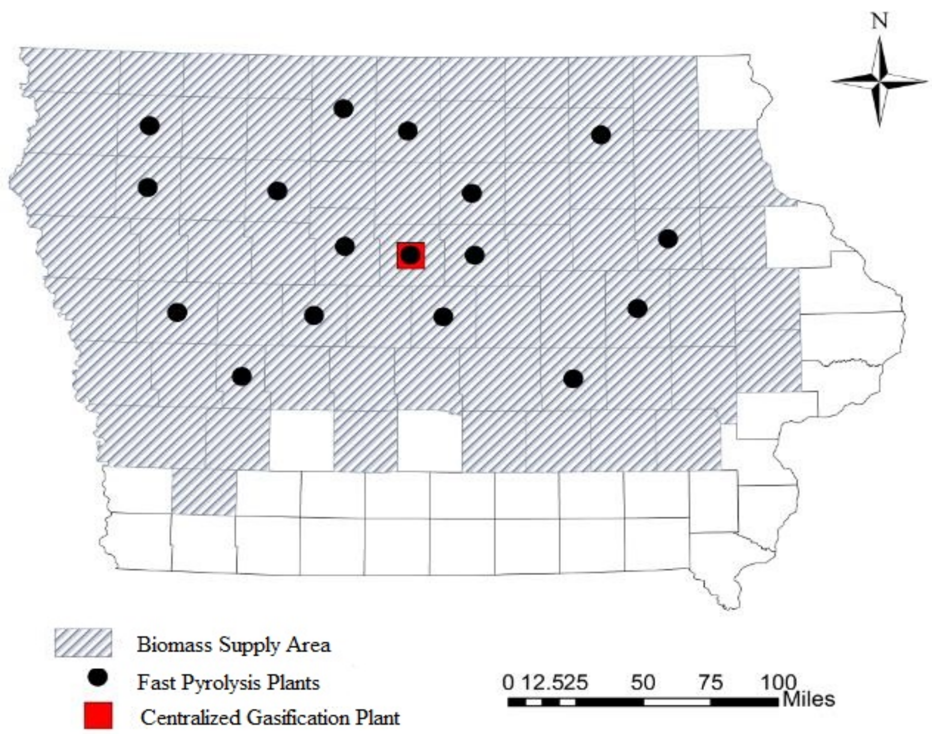


Figure 2.3 Optimal facilities locations in deterministic case

In general, feedstock production and logistics constitute more than 35% of the total production cost of advanced biofuel [57], and logistics associated with moving biomass from farmland to biorefinery can make up 50% to 75% of the feedstock cost [58] includes the annual itemized costs in deterministic case. Total shipping cost accounts for 14% of the total cost; biomass collecting cost accounts for 18% of the total cost; total capital cost accounts for about 25% of the total cost; conversion cost accounts for 43% of the total cost. In the category of shipping cost, biomass shipping cost is the most significant (54%). These results are consistent with the range reported in the literature [57, 58].

Table 2.2 Annual itemized costs in deterministic case (million dollars)

Biomass collecting cost	416.93
Total capital cost	604.33
Capital cost of the centralized refining facility	184.06
Capital cost of the fast pyrolysis facility	420.27
Total shipping cost	334.04
Biomass shipping cost	181.99
Bio-oil shipping cost	146.80
Biofuel shipping cost	5.25
Conversion cost	1020.20
Total	2375.51

2.4.3 Analysis for stochastic case

The uncertainties under consideration include biomass availability, technology advancements, and biofuel price. Technology advancement uncertainty is represented by the probabilistic distribution of two conversion ratios. Historical data for corn stover yield and retail gasoline prices are available to estimate the distributions. In this case study, moment matching method has been employed to generate the probabilistic scenarios. Statistics such as mean, variance, skewness, and kurtosis are used for moment matching.

This non-linear optimization problem is solved by applying a heuristic of changing an initiating value until a satisfactory solution is obtained. The General Algebraic Modeling System (GAMS) is utilized to solve the moment matching problem, and a scenario tree with a size of 16 is generated. A summary of scenarios in the stochastic model is included in Table 2.3.

Table 2.3 Scenario summary

Scenario	Probability	Corn Stover Yield (metric ton/acre)	Gasoline Prices ($/Gallon)	Conversion Ratio $\boldsymbol{\theta_1}$	Conversion Ratio $\boldsymbol{\theta_2}$
1	0.0128	2.2066	2.2035	0.4961	0.1825
2	0.0114	2.1568	2.5758	0.4476	0.1810
3	0.1269	2.9174	2.4271	0.7770	0.2197
4	0.1130	3.1437	4.5391	0.6242	0.1993
5	0.1116	2.9115	4.4923	0.6243	0.1984
6	0.1078	2.9048	3.4381	0.6253	0.1959
7	0.1092	2.6570	3.5253	0.6229	0.2097
8	0.1255	2.9986	3.2187	0.6206	0.1963
9	0.0531	2.7582	3.3948	0.6198	0.1961
10	0.0100	2.1041	2.5689	0.3952	0.1875
11	0.0288	2.7502	3.3767	0.5742	0.1917
12	0.0164	2.6637	3.2652	0.5465	0.1925
13	0.0259	2.7056	3.3314	0.5897	0.1944
14	0.0143	2.6095	3.1129	0.5376	0.1945
15	0.1231	3.1086	4.0164	0.6265	0.1950
16	0.0100	2.0942	2.8036	0.3858	0.1562

In the stochastic case, 17 distributed fast pyrolysis plants are proposed, and all of them are at the highest capacity level. This is the same as the deterministic case and indicates that the capacity levels are insensitive to uncertainties. The numbers of biomass feedstock sites (counties) involved in the stochastic case vary based on scenarios with a maximum of 79 counties. Nine scenarios (with a total probability of 0.6) need biomass supply from more than 71 counties. The optimal locations for these facilities are represented in Figure 2.4. The shaded areas are the union set of the biomass feedstock sites involved in all of the stochastic scenarios (81 counties).

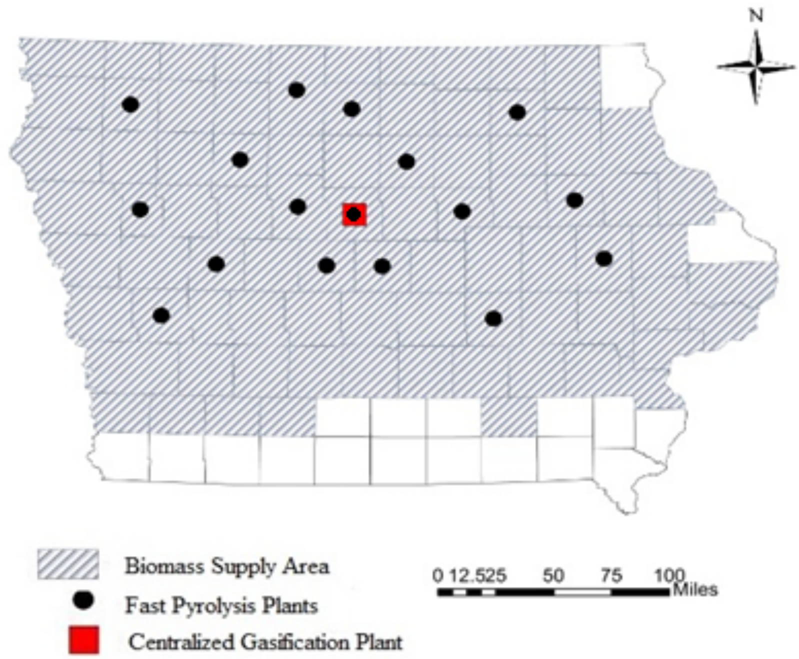


Figure 2.4 Optimal facilities locations in stochastic case

In both the deterministic and stochastic cases, 17 distributed fast pyrolysis plants are proposed but they are not at the same locations. The plants are all proposed to be built at the highest capacity level to reduce the capital cost due to the economies of scale. The centralized gasification plant will be constructed at Hamilton County in both cases, which is at the center of high corn yield counties.

Despite the similarities of both cases, differences exist for the supply chain network configurations. In the stochastic case, it is preferable to build the fast pyrolysis plants farther away from the centralized gasification and upgrading plant because biomass collection sites are more distributed due to the uncertainties in biomass feedstock supply availability. Thus, this supply chain network demonstrates the management of the trade-off between biomass availability and transportation costs.

The yearly profit in the deterministic case is 154.53 million dollars. For comparison, the numerical value of parameters used in deterministic case are the expected value of those parameters from the stochastic scenarios, thus this deterministic solution is also called the expected value solution (EV). The solution in the stochastic case is known as recourse problem solution (RP). In this case study, the yearly profit from the recourse problem is 129.57 million dollars. If we apply the decisions in deterministic case to the stochastic environment, we will get the expected yearly profit with the EV solution. This is called expected results of EV solution (EEV), which is 129.11 million dollars in this case study. The value of the stochastic solution (VSS) could be defined as $VSS = EEV - RP$. The VSS is about 0.46 million dollars, which is the direct economic benefit of considering uncertainties in the decision making process.

2.4.4 Discussion on the impact of farmers' participation

Although significant literature has investigated the environmental consequences of biomass collection from the field, limited studies have taken the social factors such as farmers' willingness to participate into consideration. However, the farmers' willingness to participate makes a direct impact on the biomass feedstock availability. Recently, an Iowa farmer survey conducted by Tyndall et al. shows that only 17% of farmers in Iowa show interest in harvesting their stover and about 37% are undecided [36]. These results suggest that about half of the farmers will not collect the corn stover in the near future. In the base case scenario, the availability factor is assumed to be 0.4, and the influence of this availability factor on the supply chain design is also investigated in this study. In this section, we discuss the impact of farmers' participation, which is represented as the availability factor δ in the model, on the decisions in both the deterministic case and stochastic case.

For the deterministic case, if the availability factor δ is less than 0.23, which means no more than 23% of the farmers would participate in corn stover collection in each county, the objective function value is equal to zero. In this case, this biofuel supply chain system is not profitable and it is optimal not to construct any facilities. When the availability factor δ is in the range of 0.23 to 0.36, the system is profitable but it could not satisfy the biofuel target of the entire state. Recall that the goal is to satisfy at least 30% of the gasoline consumption in Iowa, which is about 517 million GGE per year. Thus, at least 33000 metric ton dry biomass per day is needed at distributed fast pyrolysis plants. The biofuel supply target will be met if the availability factor δ is larger than 0.36.

Table 2.4 provides the annual itemized costs and profit for a variety of availability factors. The total capital cost, biomass collection cost, and total shipping cost increase when the availability factor δ increases from 0.3 to 0.4. This is because of the increase of the facilities' production and capacities. It should be noted that when the biofuel production capacity can meet the target biofuel demand, the total shipping cost and biomass collection cost will decrease as the availability factor increases. After that, the total capital cost will not change since the same number and capacities of facilities are planned. As a result, the yearly profit will increase as the availability factor increases. In summary, the system cost will decrease and yearly profit will increase with increase in the farmers' participation because there is more flexibility in choosing the biomass suppliers and better decisions can be reached.

Table 2.4 Annual itemized costs and profit for different δ (million dollars)

δ	0.3	0.4	0.5	0.6	0.7
Profit	69.246	154.53	200.92	232.09	256.43
Total capital cost	530.21	604.39	604.39	604.39	604.39
Biomass collecting cost	347.72	416.93	409.46	402.17	398.69
Total shipping cost	296.27	334.04	295.13	271.24	250.38
Conversion cost	840.14	1020.20	1020.20	1020.20	1020.20

Comparing Figure 2.5 to

Figure 2.3, it is observed that the locations of fast pyrolysis plants are more centralized when availability factor δ is equal to 0.7 and we only need 40 counties (rather than 71 when δ is equal to 0.4) to supply the biomass. These results not only illustrate the phenomenon that the locations of fast pyrolysis plants are sensitive to uncertainties, but also suggest that the optimal supply chain decisions will be improved by increasing biomass availability due to the additional flexibilities in choosing the biomass harvesting sites and will consequently reduce total system cost [17, 31].

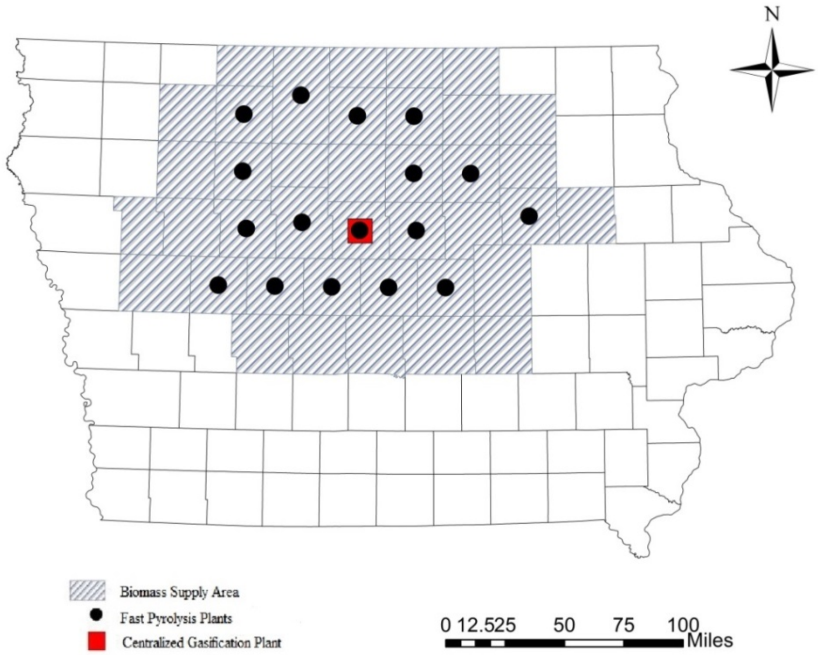


Figure 2.5 Optimal facilities locations in deterministic case (δ=0.7)

Table 2.5 shows the value of the stochastic solution (VSS) will decrease as the availability factor increase. The VSS will reduce to zero when the availability factor is larger than 0.5. It can be observed from the model that as farmers' participation increase in Iowa, the supply chain design and optimization model will become more robust. On the other hand, since the advanced biofuel industry is still at its infancy, the farmers' participation is currently at a relatively low level. Therefore, it is beneficial to apply stochastic programming framework to deal with the uncertainties and improve the decision making. This analysis provides the decision makers another insight to improve system resiliency by increasing farmers' participation.

Table 2.5 VSS for different δ

$\boldsymbol{\delta}$	EV	RP	EEV	VSS
0.3	69.25	56.25	55.74	0.51
0.4	154.53	129.57	129.11	0.46
0.5	200.92	171.82	171.76	0.06
0.6	232.09	200.93	200.93	0
0.7	256.43	222.74	222.74	0

2.5 Conclusion

Cellulosic biofuels play an increasingly important role in meeting RFS2 and reducing energy dependence. The hybrid thermochemical production pathway of bio-oil gasification which combines fast pyrolysis and gasification is one of the promising production pathways for advanced biofuel production. In this production pathway, widely distributed small-scale fast pyrolysis processing plants could avoid transporting bulky solid biomass over a long distance and a centralized gasification and fuel synthesis facility could take advantage of the economies of scale. Due to the significance of supply chain related system costs, the design of biofuel supply chain networks plays an essential role in the commercialization process.

This paper provides a mathematical programming framework with a two-stage stochastic programming approach to deal with the uncertainties in the biofuel industry. The first stage makes capital investment decisions including the locations and capacities of facilities while the second stage determines the biomass and biofuels flow. This model is a generic method for handling uncertainties in a supply chain and can be easily adapted to deal with other uncertainties and be applied to other supply chain problems. The optimization model provides methodological suggestions for decision makers of capital investment decisions and logistic decisions in the stochastic environment.

A case study of Iowa is presented to illustrate and validate this supply chain design and optimization model. The results show that uncertain factors such as biomass availability, technology advancement, and biofuel price can be pivotal in this supply chain design and optimization. The locations of fast pyrolysis plants and logistic decisions are sensitive to uncertainties while the capacity levels are insensitive. In addition, farmers' participation has a significant impact on the decision making process. It is appropriate and necessary to apply a stochastic programming framework to deal with the uncertainties, especially at a low farmers' participation level. As farmers' participation increases, the supply chain design and optimization model will become more profitable and more robust against the uncertainties along the supply chain.

In summary, this paper provides a modeling framework to study the advanced biofuel production pathway under uncertainty. Our study is subject to a number of limitations. Firstly, we assume the sustainability factor and farmers' participation are the same for each county. However, these factors may vary based on the land characteristics and agricultural management practices. Additional constraints such as water use constraints [59] can be included to better describe biomass availability. Secondly, we assume the transportation cost within counties is negligible, which could impact the supply chain design and decision making. Thirdly, we consider three sources of uncertainties and more uncertainty factors can be considered. Last but not least, only one set of scenarios is generated in this paper; more scenarios could be generated to test the stability of the stochastic results. We shall address these limitations in our future research.

CHAPTER 3 TECHNO-ECONOMIC ANALYSIS OF BIOFUEL PRODUCTION CONSIDERING LOGISTIC CONFIGURATIONS[2]

Abstract

In the study, a techno-economic analysis method considering logistic configurations is proposed. The economic feasibility of a low temperature biomass gasification pathway and an integrated pathway with fast pyrolysis and bio-oil gasification are evaluated and compared with the proposed method in Iowa. The results show that both pathways are profitable, biomass gasification pathway could achieve an Internal Rate of Return (IRR) of 10.00% by building a single biorefinery and integrated bio-oil gasification pathway could achieve an IRR of 3.32% by applying decentralized supply chain structure. A Monte-Carlo simulation considering interactions among parameters is also proposed and conducted, which indicates that both pathways are at high risk currently.

3.1 Introduction

As a renewable substitute for petroleum fuels, biofuels have attracted increasing attention for economic, environmental, and energy security considerations. First-generation biofuels could be relatively easily converted to transportation fuels but lead to food versus fuel dilemma. Cellulosic biofuel feedstock such as corn stover, switchgrass, and woody biomass does not compete with food supply but highly recalcitrant [2]. US

---

[2] This chapter of dissertation has been published in Bioresource Technology

Environmental Protection Agency (EPA) revised the Renewable Fuel Standard in 2007, which aims to accelerate the domestic biofuel production and consumption. The Revised Renewable Fuel Standard (RFS2) mandates that by the year 2022, at least 16 billion gallons per year of cellulosic biofuels will be produced and consumed in the US [6]. However, cellulosic biofuel production has been significantly below the blending targets established by the RFS2 due to technical immaturity and feedstock availability issues [8].

Lignocellosic biomass could be converted into bio-oil via pyrolysis, and the biomass pyrolysis can be followed by bio-oil cracking, gasification, or hydroprocessing to produce transportation fuels [60]. The mechanism research shows that fast pyrolysis of cellulose biomass yields to products such as pyrans, furans, and linear small molecular compounds [61]. The pyrolysis behaviors and structural features are significantly affected by the process conditions [62]. Researchers also use thermogravimetric analysis coupled to Fourier transform infrared spectroscopy to analysis the evolution of typical pyrolysis products [63].

The major challenge faced by the cellulosic biofuel industry is that investors are not willing to take the risk to construct commercial scale facilities, and lack of real facility cost information for the production systems prohibit the improvement of production system to reduce costs and uncertainty [8]. Techno-economic analysis (TEA) has been widely adopted to overcome this dilemma. Process models are developed to simulate the production systems at a commercial scale. Materials and energy balances are developed. Cost analysis is then employed to evaluate the economic feasibility of the production system at commercial scale [64]. TEA, as a simulation approach, is highly dependent on the model assumptions, which could lead to significant inaccuracy and even errors.

Another major barrier for commercialization of cellulosic biofuels is transporting bulky solid biomass over a long distance. This is mainly caused by the low energy density of lignocellulose biomass and a large collection radius due to the limitation of biomass availability. In general, logistics cost for transport biomass from farmland to biorefinery can make up 50% to 75% of the feedstock cost [65] and more than 35% of the total production cost of advanced biofuel is feedstock cost [57].

TEA studies typically focus on the technical and economic performance for a single facility and neglect the upstream biomass collection and transportation as well as the downstream biofuel transportation and distribution. However, with the importance of supply chain configurations in the economic feasibility evaluation of cellulosic biofuels, TEA should incorporate the supply chain configurations explicitly rather than the simplify assumption of a flat feedstock cost and biofuel price at the facility gate. Recently, researchers have worked on incorporating pre-determined simple supply chain configurations to estimate biomass feedstock cost for integrated pathways [66, 67]. However, in reality, feedstock availability, logistic cost, biofuel demands will all affect evaluation of economic feasibility [68]. This serves as the major motivation for this proposed approach to incorporate logistic settings into the techno-economic analysis.

There has been an increasing body of literature on supply chain network design for the biofuel industry [17, 68, 69]. Design and management of logistic flow includes the raw materials, work-in-process, and finished products from source of raw materials to the point of consumption [26]. In order to incorporate supply chain design into TEA study, logistic information such as biomass availability, transportation cost, and demands distributions is necessary. A decision method and optimization model is necessary to determine the optimal facility locations and capacities as well as the logistic flow decisions for biomass supply and biofuel distribution.

The remainder of the paper is organized as follows: in Section 2, the proposed TEA method with logistic settings is introduced. In Section 3, we illustrate the method with a case study of comparing two competitive pathways in Iowa, namely low temperature biomass gasification pathway and fast pyrolysis (FP) integrated with bio-oil gasification pathway. Finally, the paper concludes with a summary of research findings in Section 4.

3.2 Materials and Methods

In this section, the proposed TEA method with logistic configurations is introduced. Materials and conversion pathways are chosen based on current technology and feedstock availability. Methods for technical and economic analysis are discussed.

3.2.1 TEA method with logistic settings

This proposed TEA method with logistic configurations contains three main steps: cost estimations based on traditional standalone TEA, design and evaluation of supply chain configurations, and economic feasibility assessment under realistic supply chain. In the first step, the investment for a single facility based on a traditional standalone TEA literature is evaluated. An assessment on the relationship between plant sizes and economic performance based on the rules of economies of scale and time value of money provide candidate plant sizes for the supply chain design. The second step is to design the biofuel supply chain configurations based on the conversion pathway and feedstock availability of the region under assessment. Mathematical models are formulated to provide the decision support for the supply chain design. The third step includes economic performance assessment considering the logistic configurations, and risk assessment with Monte-Carlo simulation.

The motivation of this proposed TEA method is to introduce supply chain design into traditional TEA to achieve a more comprehensive analysis and realistic economic assessment results. In the conventional TEA which has been commonly used in the literature, assumptions such as flat feedstock cost and biofuel price at the facility gate have been adopted [9, 11, 53]. These assumptions have received significate concerns. In this proposed TEA with logistic considerations, no uniform feedstock prices at facility gates are assumed. Instead, feedstock cost is estimated by the farm gate collection cost and the shipping cost from farm to facility. The feedstock collection costs vary due to collection methods and quantities. A regression analysis is employed to estimate collection cost [45]. The feedstock (crop residuals, such as corn stover) availability is assumed to be proportional to the crop yield. The feedstock shipping cost, the intermediate product shipping cost, and the biofuel shipping cost are all assumed to be proportional to the shipping distance. Biofuel market prices are based on U.S. Energy Information

Administration (EIA) projection, and the biofuel demand amounts and locations are based on the population distribution in the geographic region.

3.2.2 Materials and technologies

To illustrate the proposed TEA method, a case study based on Iowa is conducted. Corn stover has been chosen as the cellulosic biomass feedstock in this study due to its abundance in Iowa (Wilcke and Wyatt, 2002). The final biofuel product is assumed to a drop-in fuel which is ready for vehicle consumption.

In this study, we have thus chosen integrated pathway with fast pyrolysis and bio-oil gasification to illustrate the proposed TEA method. It has been suggested that hybrid pathways, such as integrating fast pyrolysis and downstream upgrading process such as gasification would be a viable option for commercial scale. This is due to the flexibility to accommodate a decentralized supply chain structure and also advantage of economies of scale (Li et al., 2015; Manganaro and Lawal, 2012).

On the other hand, low temperature (870℃) biomass gasification pathway has typically been brought up for comparison with the integrated bio-oil gasification pathway. Therefore, the conversion pathways under consideration in this study include low temperature biomass gasification pathway and hybrid fast pyrolysis with bio-oil gasification pathway. BMG (biomass gasification) and BOG (bio-oil gasification) are used as abbreviations for these two pathways in the following sections.

3.2.3 Technical analysis

The conversion process models and mass and energy balance information are based on the existing literature (Li et al., 2015; Swanson et al., 2010). Figure 3.1 shows the process flow diagrams for biomass gasification pathway and integrated bio-oil gasification pathway. The main assumptions such as capital cost estimation, plant size, target IRR, and facility life are the same in both studies, while the balance of plant (BOP) and annual operating hour rate are in similar range which is typical in TEA studies (11% and 0.85 in

BMG TEA while 12% and 0.9 in BOG TEA) [10, 53]. These assumptions are preserved in this study to reflect the similarity as well as slight differences between these two pathways.

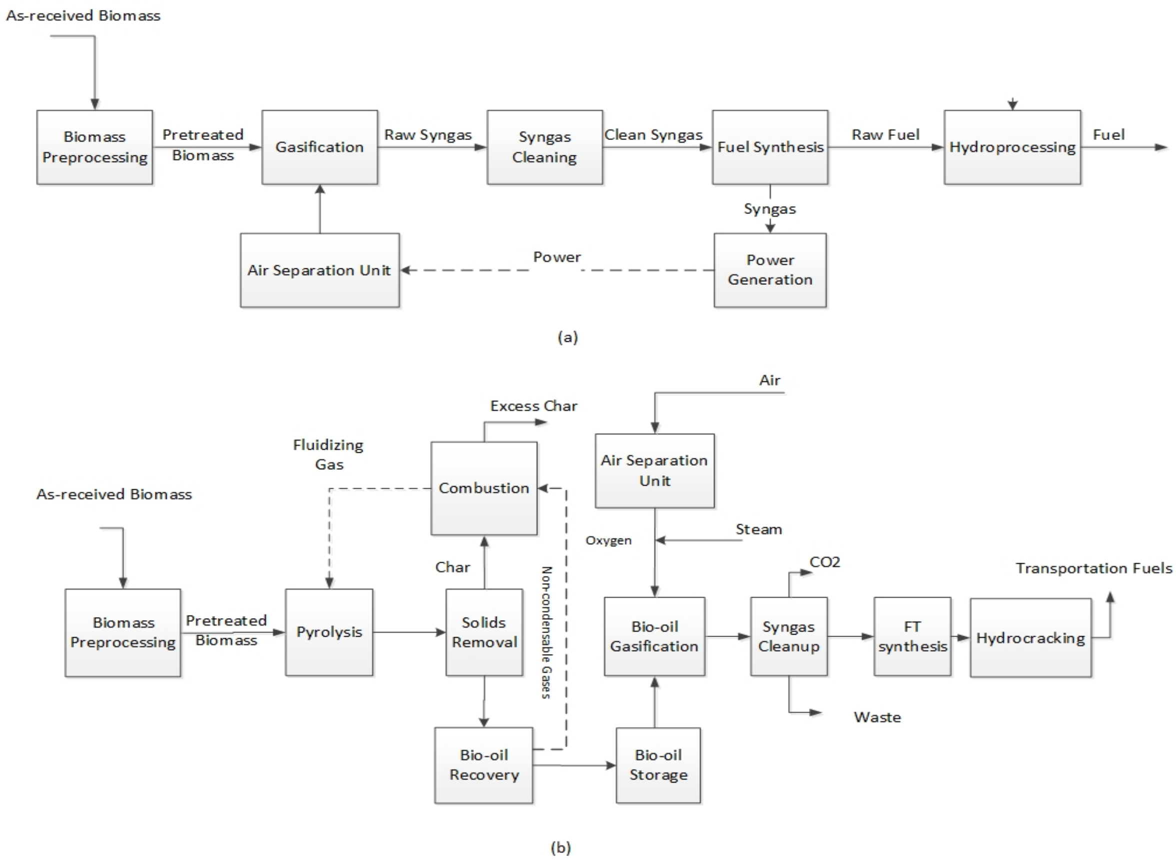


Figure 3.1 Process flow diagrams for biomass gasification pathway (a) and bio-oil gasification pathway (b)

3.2.4 Economic analysis

An *n*th plant scenario with facility life of 20 years is considered in this study. Estimations for capital investment, installed equipment cost, and annual operating cost are based on the literature (Li et al., 2015; Swanson et al., 2010). The capital and operating costs for different capacities are estimated based on the rules of economies of scale and time value of money. The economies of scale refers to the concept that the cost per unit of output would generally decrease with the increasing scale of a facility [70]. Eq. (1) is adopted to estimate costs from the base costs.

$$cost_{new} = \left(\frac{I}{I_0}\right) * cost_0 * \left(\frac{size_{new}}{size_0}\right)^n \quad (1)$$

$cost_0$ is the base equipment cost, $size_0$ is the size of base equipment, I_0 is the inflation index of the base year. $cost_{new}$ is the new equipment cost, $size_{new}$ is the size of new equipment, I is the inflation index of the calculated year. n is the scaling factor with a typical range from 0.6 to 0.8. "Sixth-tenth rule" is typically adopted in TEA studies[71]. However, since BMG and BOG pathways are both relatively immature, a more conservative scaling factor of 0.7 is used in this study [66]. All monetary figures have been adjusted to 2013 dollars based on inflation. The income tax rate is assumed to be 39%. The biofuel market price is assumed to be 3.5 $/GGE, which is the average gasoline projection price by EIA for the next twenty years [72].

3.3 Results and Discussion

In this section, the results for the case study with the proposed method are discussed.

3.3.1 Cost estimations for supply chain design

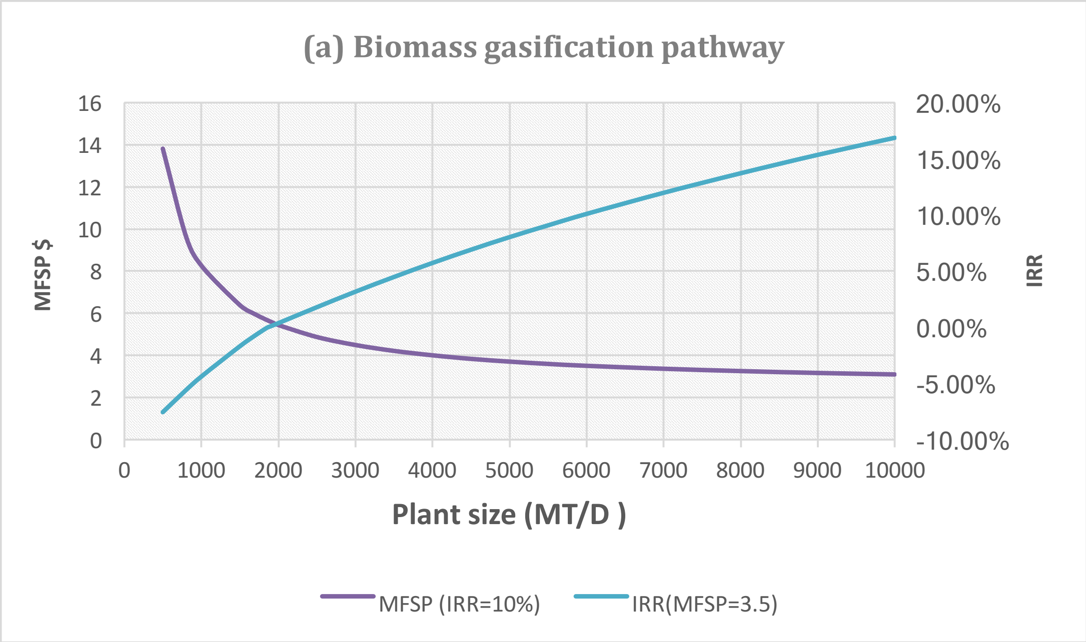
In the supply chain design and analysis, the main results of economic performance for a single facility from conventional TEA are adopted. Candidate plant sizes are chosen based on a single facility economic performance analysis at different capacities. Capital cost and operation cost at candidate plant sizes are also evaluated.

For a 2000 metric ton per day (MT/D) plant, fuel yield for BMG is about 293 MT/D, while the fuel yield for BOG is 239 MT/D. The total capital investment (TCI) for BMG is 559.9 million dollars and TCI for BOG is 510 million dollars. The minimum fuel selling prices (MFSPs) for these pathways are 5.43 $/GGE and 5.59 $/GGE respectively. These high MFSPs reflect the current status of technology. Table 3.1 summarizes the breakdown of the capital costs and operating costs. These results indicate that for a single plant size of 2000 MT/D, the BOG has a lower TCI while BMG pathway has a higher fuel yield. As a result, BOG and BMG pathways perform similarly in terms of MFSP.

Table 3.1 Breakdown capital costs and operating costs

Pathway	BMG	BOG
Installed cost breakdown for *n*th plant (million dollar)		
Preprocessing	25.5	25.4
Gasification	31.7	22.6
Syngas cleaning	32.9	23.2
Fuel synthesis	66.0	47.4
Hydroprocessing	33.1	25.1
Power generation	43.7	NA
Fast Pyrolysis & Combustion	NA	84.4
Air separation unit	21.9	15.4
Balance of plant	30.6	29.2
Total installed cost	285.3	273.0
*n*th plant results (million dollar)		
Indirect cost	120.4	92.0
Fixed capital investment	486.8	444.0
TCI	559.9	510.0
Annual operating cost for *n*th plant (million dollar)		
Fixed costs	13.5	13.4
Variable costs	14.6	9.5
Feedstock	57.3	54.3
Capital depreciation	24.7	21.9
Average income tax	20.2	17.9

It should be noted that multiple facility sizes are considered in the biofuel supply chain design and configuration phase. Figure 3.2(a) illustrates the changes in facility IRR and MFSP at different facility sizes of BMG. Similar trend for BOG can be observed as shown in Figure 3.2(b) [11]. If we assume the MFSP to be 3.5 \$/GGE, the relationship between the facility IRR and capacity could be analyzed. For BMG pathway, any facility capacity larger than 1900 MT/D would give a positive IRR and a 10% IRR is achieved at 6000 MT/D. These results indicate that for both pathways, plant size of 2000 MT/D is near the break-even point, and a larger plant size is favorable due to the economies of scale. The candidate plant sizes have been chosen to be 2000 MT/D, 4000 MT/D, and 6000 MT/D.



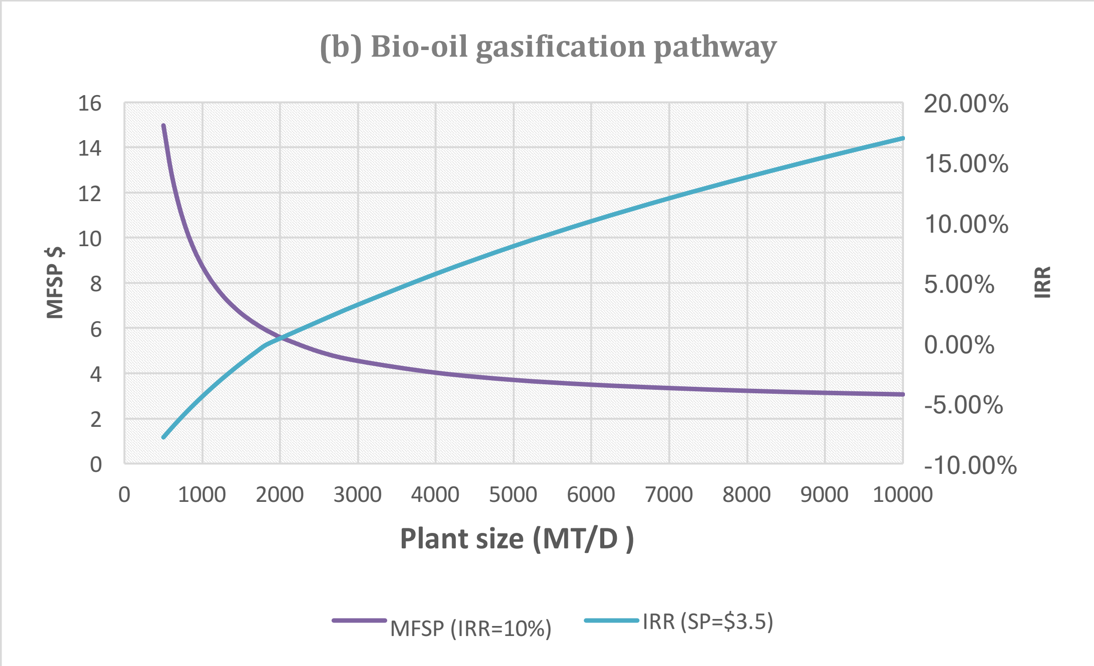


Figure 3.2 Variation of MFSP and IRR with plant sizes for biomass gasification pathway (a) and bio-oil gasification pathway (b)

According to RFS2, at least 36 billion gallons per year of renewable fuels will be produced by 2022, which is approximately 28% of the national gasoline consumption. In this study, the biofuel demand for Iowa is set to be 550 million GGE per year, which could satisfy about 30% of the gasoline consumption in Iowa.

The capital and operating costs for base size of facility are based on existing literature [9, 11, 53]. Key parameters used in supply chain design model such as annual operation costs and capital investment costs are summarized in Table 3.2. Note that the annual operation costs only include fixed operation costs and variable operation costs. The operation costs do not include income tax, feedstock cost, and capital depreciation because these costs would be included in the supply chain model.

Table 3.2 Key economic parameters at different plant sizes

Scenario	Plant size	Annual fuel output (million GGE per year)	Capital cost (million dollar)	Annual operation cost (million dollar)
BMG	2000	32	560	32
BMG	4000	65	909	58
BMG	6000	97	1208	82
BMG	34000	550	4069	397
FP	2000	NA	63	26
FP	4000	NA	102	36
FP	6000	NA	136	65
BOG	42000	550	3398	254
BOG＋FP	42000	550	4300	471

3.3.2 Supply chain design

The supply chain system schematics for BMG and BOG pathways are shown in Figure 3.3. Biomass is firstly collected at county level, and transported to the biorefinery facility (facilities) to produce transportation fuels. The biofuels are then distributed to demand areas which are based on metropolitan statistical areas. It is assumed that each biomass feedstock location can support multiple facilities, and that each facility can acquire feedstock from multiple biomass supply locations. The locations for all biomass feedstock sites, biorefineries, and demand sites are assumed to be the centroids of counties. Mixed integer linear programming models are employed to decide the optimal plant sizes and locations of the facilities to maximize the profit. The mathematical models are adapted from literature [68]. The General Algebraic Modeling System (GAMS) software is used to obtain numerical results.

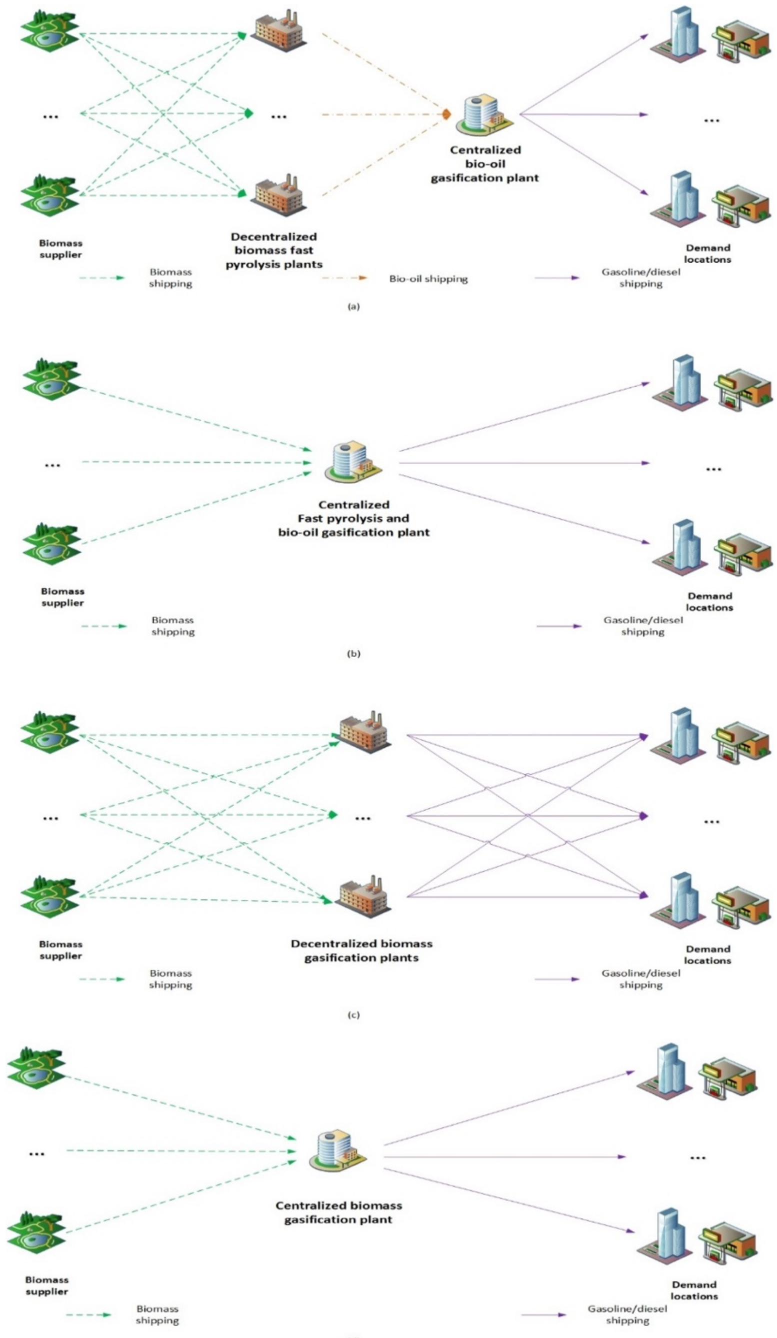


Figure 3.3 Supply chain configurations for each scenario

Four scenarios are considered in this study, two for BOG pathway and two for BMG pathway. In the first scenario, the supply chain contains one centralized 42000 MT/D BOG facility with multiple decentralized fast pyrolysis facilities with variable plant sizes converting biomass to bio-oil (Figure 3.3(a)). In the second scenario, only one integrated 42000 MT/D facility with both bio-oil gasification and fast pyrolysis at the same site (Figure 3.3(b)). In the third scenario, multiple BMG facilities at different plant sizes would be constructed to satisfy the biofuel demand (Figure 3.3(c)). In the fourth scenario, only one centralized 34000 MT/D BMG facility will be constructed to satisfy the demand (Figure 3.3(d)). The geographical locations and sizes of the facilities are determined optimally with the mathematical models.

The supply chain configuration designs are based on existing literature [68] , industrial relevancy, and comparison consistency. The plant sizes in the third scenario with multiple biomass gasification facilities are set to be chosen from 2000 MT/D, 4000 MT/D, and 6000 MT/D based on the analysis in section 3.1. The plant sizes of decentralized fast pyrolysis plants in the first scenario are also set to be chosen from 2000 MT/D, 4000 MT/D, and 6000 MT/D for consistent comparison.

3.3.3 Numeral results

The computational results of each scenario are presented and discussed in this section. First scenario achieves an IRR of 3.32% with a total project investment of 4,481 million dollars. Second scenario yields an IRR of 2.21% with a total project investment of

4,300 million dollars. Third scenario would have an IRR of 1.80% with a total project investment of 6,949 million dollars. Fourth scenario would have an IRR of 10.00% with a total project investment of 4,069 million dollars. The fourth scenario achieves the highest IRR and requires the smallest investment.

Table 3.3 provides the annual itemized costs for each scenario. The results demonstrate the trade-off between feedstock shipping cost and the capital investment under a variety of supply chain configurations. By adding 181 million dollars capital investment, the first scenario could increase the IRR by 1.1% comparing to the second scenario. These additional capital investments will lead to a saving of over 200 million dollars on shipping cost per year. This’s mainly because the decentralized fast pyrolysis facilities could decrease the biomass shipping cost significantly without significantly increasing capital investment [73]. These results indicate that BOG pathway is more suitable for decentralized supply chain structure.

Table 3.3 Annual itemized costs (million dollars)

Scenario	1 BOG (D)	2 BOG (C)	3 BMG (D)	4 BMG (C)
Biomass collection cost	445	475	378	380
Biomass shipping cost	181	411	169	305
Bio-oil shipping cost	16	NA	NA	NA
Biofuel shipping cost	6	6	4	6
Operation cost	631	471	469	397
Capital depreciation	195	187	302	177
Income tax	185	171	248	271
Total	1660	1710	1570	1530

On the other hand, comparing between the third and fourth scenario, although the third scenario could save about 148 million dollars on shipping cost per year, additional 135 million dollars costs on capital depreciation and 72 million dollars on operation cost per year would incur. In other words, the saving in shipping cost could not offset the increase in capital investment and annual operation costs. These results show that the effect of economies of scale is more significant than decentralized supply chain to decrease the logistic costs for BMG pathway. This is due to the high capital investment of this biofuel production pathway. Therefore, larger facility is preferred for BMG pathway even through it cost would have more feedstock shipping cost.

Comparing between the two pathways, i.e., compare the first scenario with the third scenario, and compare the second scenario with the fourth scenario, the biomass collection cost as well as biomass shipping cost are higher for BOG pathway under both supply chain structures because more biomass feedstock is needed in BOG pathway due to lower fuel conversion yield. In the meantime, the operation cost is higher for BOG pathway under both supply chain structures than BMG pathway, which is due to the process complexity. These results indicate that at the current stage of technology, BMG pathway has higher efficiency of energy conversion than BOG pathway. In addition, these comparisons show that different production pathways could have its preferred supply chain structure.

Decentralized supply chain structure is more suitable for BOG pathway, while centralized supply chain structure works better with BMG pathway.

Figure 3.4 shows the facility locations and plant sizes in each scenario. The darkness in county indicates the corn stover yield, darker color means higher yield. The shaded areas represent the countries that provide the biomass feedstock, and star symbols represent the locations of biofuel demand. The black dots represent the locations of the decentralized facilities and the black square represents the location of the centralized facility. 42 counties will support ten 4000 MT/D and one 2000 MT/D decentralized facilities in the first scenario, and the centralized facility is located in Hamilton County (Figure 3.4(a)). 47 counties will support the centralized facility in the second scenario, which is also located in Hamilton County (Figure 3.4(b)). 38 counties will support five 6000 MT/D and two 2000 MT/D decentralized facilities in the third scenario (Figure 3.4(c)). 37 counties will support the centralized facility located in Wright County, which is next to Hamilton County (Figure 3.4(d)).

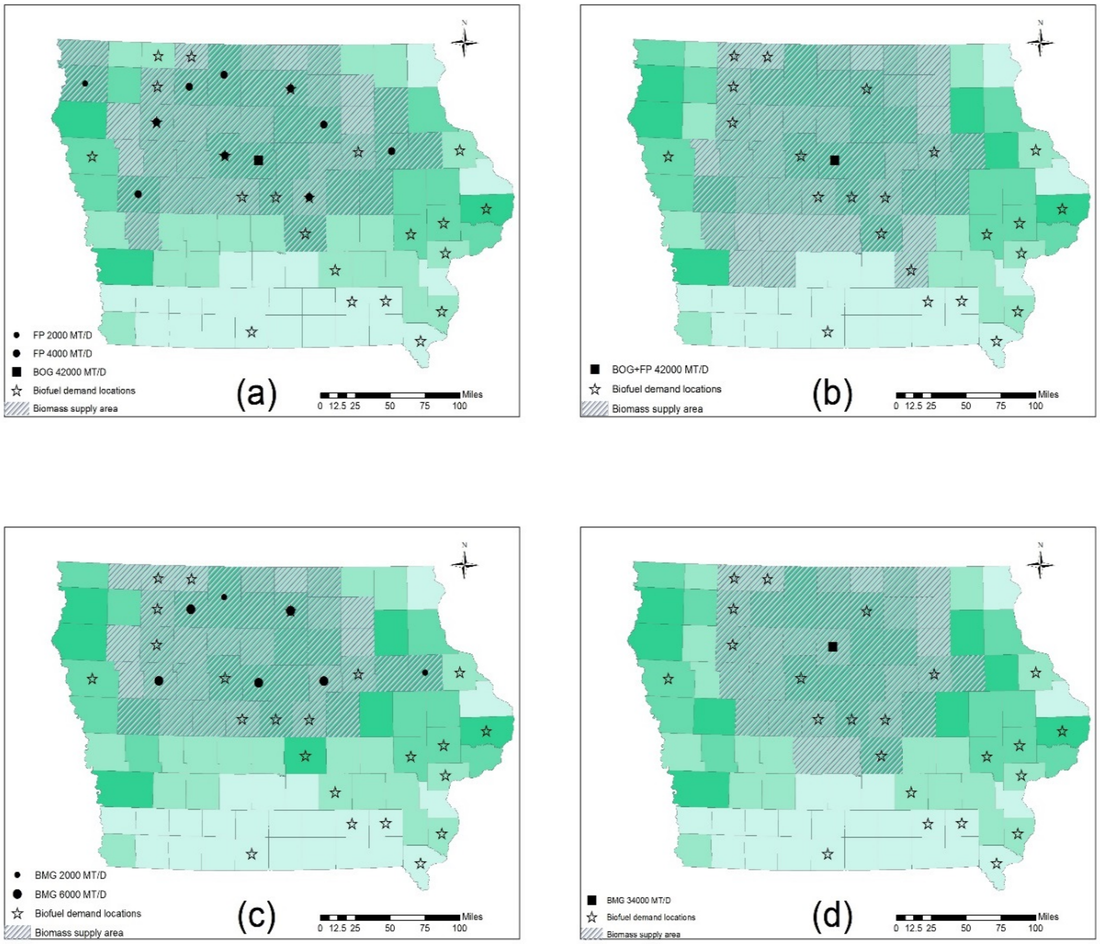


Figure 3.4 Optimal facility locations in each scenario

All scenarios prefer to build facilities in the northern part of Iowa which has higher corn yield, thus higher biomass availability. The majority of the decentralized facilities in the first scenario are built at the moderate plant size (4000 MT/D) while the majority of the decentralized facilities in the third scenario are built at the highest plant size (6000 MT/D). This is because the effect of economies of scale is more significant for BMG pathway. Centralized supply chain configuration needs more involved counties than decentralized supply chain configuration for both pathways. Because decentralized facilities could collect biomass from nearby counties which has higher biomass availability. On the other hand, BOG pathway needs more biomass feedstock than BMG pathway for both supply chain configurations due to its lower conversion efficiency. As a result, the supply chain of BOG pathway involves more countries as feedstock suppliers and the locations of decentralized facilities are more distributed.

3.3.4 Discussions

Typically, point estimators of MFSPs and IRR are provided as the economic assessment results. As a simulation method, the results of TEA are highly dependent on the assumptions. Sensitivity analyses have been adopted as a paradigm to evaluate uncertainties in the parameters. The sensitivity analyses in traditional TEA adjust a single parameter at a time and evaluate the impact on MFSP, IRR or net present value (NPV). This method could not account for the interactions between the parameters since those may not always be independent. Recent TEA literature also includes the Monte-Carlo simulations as a part of uncertainty analysis which can incorporate simultaneous changes for multiple parameters [10, 64]. However, the arbitrary selection of the probability distributions and assumption of independence weaken its superiority to the sensitivity analyses. This serves as the major motivation to conduct Monte-Carlo simulation considering interactions between parameters in this study. Multivariate distribution of key factors can be assigned when the correlations of the parameters are not negligible. Monte-Carlo simulation of BMG pathway for a single 2000 MT/D facility is conducted in this section to illustrate this method.

Based on the sensitivity results of BMG pathway, TCI, feedstock cost, and compressor install factor have the most significant impact on MFSP [9]. In this Monte-Carlo simulation, IRR, TCI, feedstock cost (FC), and compressor install factor (CIF) are selected as key parameters for analysis. IRR is added for a consistent comparison with BOG pathway. The probability distributions of these parameters are chosen based on literature. 5000 Monte-Carlo simulation runs are conduced and analyzed using R software. The simulation results are used to analyze the empirical distribution of MFSP and give interval estimators of MFSP. These estimators could capture the economic feasibility of BMG pathway under uncertainty.

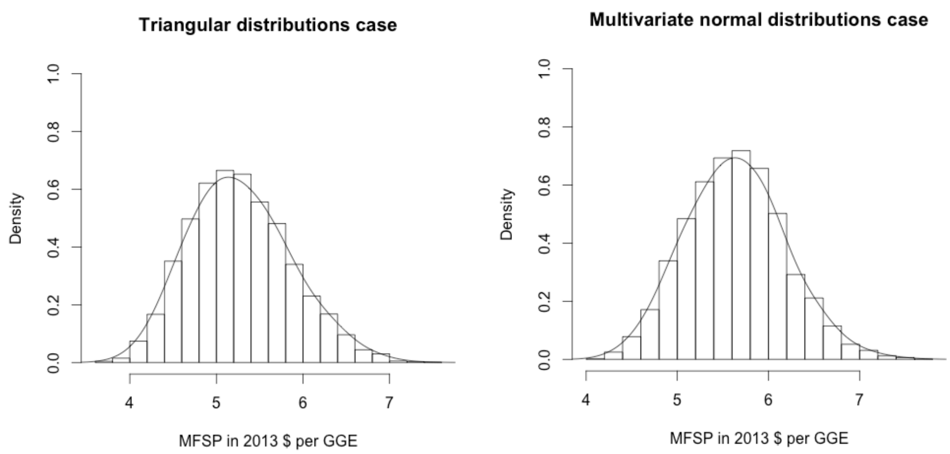
Triangular distributions and normal distributions are commonly suggested distributions for parameters in the literature [74]. Two cases are considered in this study. In the first case, all of these parameters are assumed to follow triangular distributions with the same ranges used in the sensitivity analysis due to data availability limitation. In the second case, these variables are assumed to follow multivariate normal distribution with means vector $\underline{\mu}$ equal to their base level and variances equal to one sixth of the range of their triangular distributions. Analysis based on similar settings for BOG pathways are available in literature [11]. It shows that the normal distributions case has a larger mean for MFSP (5.46 \$/GGE to 6.23 \$/GGE). Both cases show that more than 66% of runs have MFSP exceeding 5 \$/GGE. Note that these results are based on the independent assumptions among parameters.

Furthermore, in order to capture the interactions among parameters, the correlation coefficient of TCI and CIF ρ is assumed to be 0.5 to illustrate the positive correlations between these two parameters while other correlations are assumed to be zero to indicate independence. In other words, these four parameters follow multivariate normal distribution in Eq. (2).

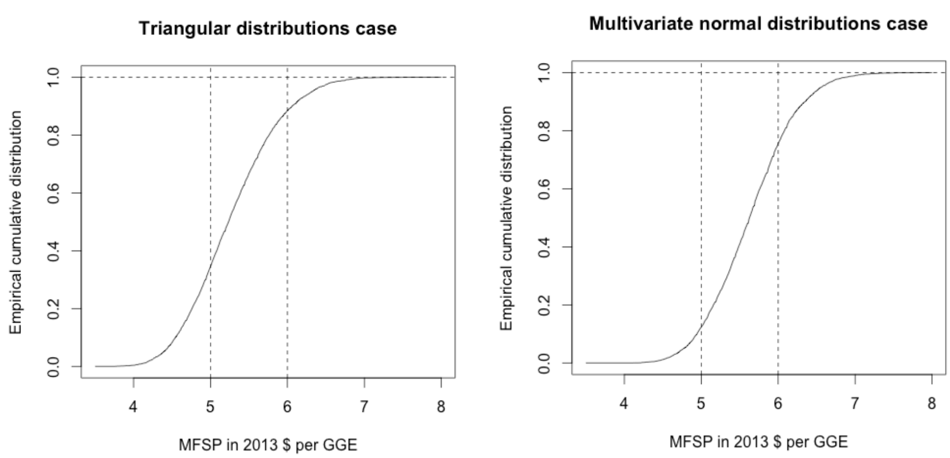
$$N\left(\underline{\mu}, \Sigma\right), \ \ where \ \ \underline{\mu} = (\mu_{IRR}, \mu_{FC}, \mu_{TCI}, \mu_{CIF})'$$

$$and\ \Sigma = \begin{bmatrix} \sigma_{IRR}^2 & 0 & 0 & 0 \\ 0 & \sigma_{FC}^2 & 0 & 0 \\ 0 & 0 & \sigma_{TCI}^2 & \rho\sigma_{TCI}\sigma_{CIF} \\ 0 & 0 & \rho\sigma_{TCI}\sigma_{CIF} & \sigma_{CIF}^2 \end{bmatrix} \qquad (2)$$

Figure 3.5(a) includes the probability density function of MFSP from MC simulation. The normal distributions scenario has a higher mean than the triangular distribution scenario (5.5 \$/GGE to 4.5 \$/GGE) and this is consistent with the BOG literature [11]. The empirical cumulative distribution of MFSP is shown in Figure 3.5(b). These results show that 33% of the simulation runs yield a MFSP that are less than 5 \$/GGE and 15% of the runs have a MFSP that exceed 6 \$/GGE in the triangular distribution scenario. Meanwhile, in the normal distribution scenario, 12% of the simulation runs yield a MFSP that are less than 5 \$/GGE and 25% of the runs have a MFSP that exceed 6 \$/GGE.



(a)



(b)

Figure 3.5 Probability density function (a) and empirical cumulative distribution (b) of MFSP from MC simulation

Both distributions show the range of MFSP is about 4-7 \$/GGE. A 95% confidence interval based on empirical distribution is [4.2, 6.7] \$/GGE for triangular distribution scenario and [4.5, 6.7] $/GGE for normal distribution scenario. The wide range of MFSP and their high probability to exceed 6 $/GGE are consistent with the Monte-Carlo simulation results of BOG pathway. These confidence intervals also verified that both pathways seem to be economically not attractive for a 2000 MT/D facility. In addition, assumptions for distribution as well as its variance covariance structure can take significant impact on the uncertainty analysis.

The analyzes show that BMG pathway could achieve an IRR of 10.00% and BOG pathway could achieve an IRR of 3.32% by assuming the biofuel market price is 3.5 $/GGE. IRR and MFSP are indicators of economic feasibility for a pathway. Based on the literature, TEA studies of a single 2000 MT/D facility typically set IRR to be 10% and evaluate the MFSP to compare against prevailing market price. The MFSP of thermochemical cellulosic biofuel pathways ranges from 1.82 $/GGE to 7.32 $/GGE due to variations in technical settings and assumptions [8]. There are also studies use IRR as analysis output. For instance, the expected value of facility IRR is 13.1% for bio-oil upgrading pathways and 8.4% for bio-oil gasification of biohydrogen under biofuel market price [10, 64]. As indicated in Figure 3.2, single facility IRR is highly affected by plant sizes and fuel selling prices. However, an IRR higher than 10% often indicates an attractive investment. It's known that the biofuel producer not only could gain profit by selling biofuel, but also could generate revenue by selling the renewable identification numbers (RINs) credits [75]. RINs are used to demonstrate compliance with the RFS. Cellulosic RINs price was about 0.78 $/RIN in 2012 and 0.42 $/RIN in 2013 [76]. As shown in Figure

3.6, with a cellulosic RINs price of 0.5 $/RIN, BMG pathway could achieve an IRR of 16.22% and BOG pathway could achieve an IRR of 9.53%, which would make both pathways more economic competitive.

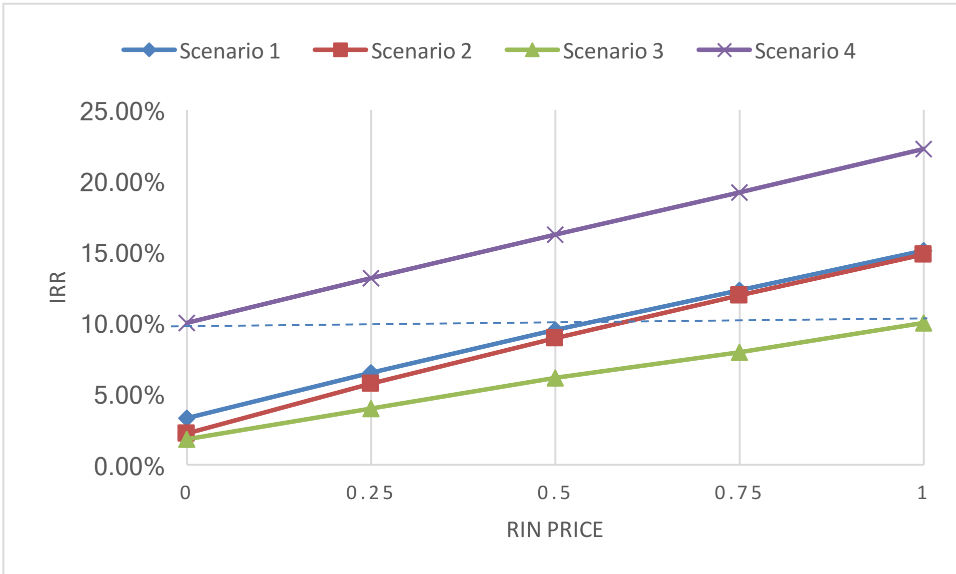


Figure 3.6 The relationship between RIN price and IRR

In this section, the results in above analysis are summarized to provide a comprehensive comparison between BOG pathway and BMG pathway.

For a single plant size of 2000 MT/D without consider the supply chain configurations, the economic performance of BOG and BMG pathways are similar in terms of MFSP. For a single facility, the break-even point of plant size is 1900 MT/D for both pathways. BOG pathway could achieve a 10% IRR at plant size of 5000 MT/D while BMG pathway could achieve 10% IRR at plant size of 6000 MT/D. Large plant size is preferred for both pathways due to the economies of scale.

The supply chain analysis results show BMG pathway is more economically feasible than BOG pathway in Iowa when realistic supply chain configurations and constraints are considered. Different production pathways could have its preferred supply chain structure. BOG pathway is more suitable for a decentralized supply chain structure while BMG pathway is more suitable for a single facility supply chain structure. The supply chain configuration demonstrates the trade-off between feedstock shipping cost and the capital investment of multiple facilities in different scenarios.

Both cases in the Monte-Carlo simulation results for single 2000 MT/D facility show that the range of MFSP is about 4-7 $/GGE for BMG pathway. These results indicate that even through BMG pathway has better economic performance than BOG pathway, both pathways are at high risk at this point.

This study is subject to a number of limitations. First, the supply chain parameters such as fuel prices, shipping costs are assumed to be deterministic, stochastic should be considered in future study. Second, the capital and operation cost at different plant sizes are roughly estimated by economies of scale. Future study could achieve a more precise estimate by modeling the diversity of parameters such as scaling factors and labor costs. It should be noted that as a general framework, other biofuel production pathways could be evaluated considering supply chain configurations using the same procedures.

3.4 Conclusion

In the paper, a new TEA method considering supply chain configurations has been introduced. The proposed approach is illustrated with a case study to compare two competitive pathways in Iowa. The results indicate that biomass gasification pathway has better economic performance than hybrid fast pyrolysis and bio-oil gasification pathway under current technology status. Hybrid fast pyrolysis and bio-oil gasification pathway is more suitable for a decentralized supply chain structure while biomass gasification pathway is more suitable for a single centralized facility supply chain structure.

CHAPTER 4 A FARM-LEVEL PRECISION LAND MANAGEMENT FRAMEWORK BASED ON INTEGER PROGRAMMING[3]

Abstract

Farmland management involves several planning and decision making tasks including seed selection and irrigation management. A farm-level precision farmland management model based on mixed integer linear programming is proposed in this study. Optimal decisions are designed for pre-season planning of crops and irrigation water allocation. The model captures the effect of size and shape of decision scale as well as special irrigation patterns. The authors illustrate the model with a case study on a farm in the state of California in the U.S. and show the model can capture the impact of precision farm management on profitability. The results show that threefold increase of annual net profit for farmers could be achieved by carefully choosing irrigation and seed selection. Although farmers could increase profits by applying precision management to seed or irrigation alone, profit increase is more significant if farmers apply precision management on seed and irrigation simultaneously. The proposed model can also serve as a risk analysis tool for farmers facing seasonal irrigation water limits as well as a quantitative tool to explore the impact of precision agriculture.

[3] This chapter of dissertation has been published in PLOS ONE

4.1 Introduction

Farmland management under climate change and population growth is a pressing challenge that is becoming increasingly important due to food security considerations. The Institute for Operations Research and the Management Sciences (INFORMS), the leading professional association in analytics and operations research along with industrial interests, encouraged researchers to address the problem of feeding millions of people throughout the world who face hunger every day. There has been a growing body of literature on crop rotations at a regional scale [12, 13], land use patterns, and policy and environment issues on a farm scale [14]. Precision agriculture has attracted increasing attention in the community of farmland management. Over the years, the precision agriculture philosophy has enriched from simply "farming by soil" to a comprehensive system including irrigation planning, phenotypic selection, vehicle guidance systems, product quality and environmental management etc. [3-5]. As the demand for agricultural products increases, water and arable land become significant factors to improved agricultural production. Each year, farmers have to make decisions about what crops to plant given knowledge about the soil on their respective farms. Farmers need to select seed and plan for irrigation carefully to ensure maximum benefit from farming. Thus, crop planning and irrigation water management on a farm scale are imperative for improved agricultural productivity and sustainable development [7].

At the farm scale, farmers have a particularly strong incentive to optimize their water usage when the irrigation water price is high and the volume of available water is limited [77]. However, optimal usage of irrigation water resources requires efficient techniques and decision making support. There are mainly two approaches for this. On one hand, seed hybrid selection is one method to improve water utilization. With the development of phenotype prediction and genotype selection, it is possible to utilize the high yield and drought resistant crop seeds. These new seed types give a farmer more flexibility to plant a variety of seeds on a farm, but also increase the difficulty for optimal pre-season seed planning. Alternatively, it is suggested that deficit irrigation is a more efficient method for water usage [78]. Deficit irrigation refers to the method that distributes a limited amount of irrigation water to satisfy essential water needs of plants [79]. Deficit irrigation could increase system benefits by saving water recourses, at the cost of individual benefits, by decreasing crop water allocation, especially during less critical periods of water demand. There are two major methods to implementing deficit irrigation for farmland. The first is to increase the interval between irrigation events. In other words, continue to irrigate with the same amount of water per irrigation as in the past but decrease the irrigation frequency (increase the number of days between irrigations). The second method of deficit irrigation is to irrigate at the same frequency as normal, but apply less water at each irrigation so that only a partial saturation level is achieved [80].

Mathematical programming has been widely used in farmland management, especially in irrigation management. Singh reviewed the literature in modeling, planning, and optimization of irrigation management with a focus on applications of different modeling techniques [81]. Sethi et al. developed a linear programming optimization model to find maximum annual net return for cropping and groundwater management [7]. The model was applied to a coastal river basin in India under different soil types, cropping patterns, and types of crops. Georgiou and Papamichail used simulated annealing and a gradient descent algorithm for reservoir and crop planning optimization [79]. Their method accounted for variable reservoir inflows and climate variability for crop planning. Wardlaw and Bhaktikul applied a genetic algorithm to optimize the delivery of water flows to minimize the distribution losses of an open race irrigation distribution system [82]. The major constraints in this study related to in-field soil moisture balances as well as canal capacities. Nagesh Kumar et al. used genetic algorithms for real-time reservoir operation management of multiple crops [83]. The study aimed to maximize the total yields from all crops considering reservoir inflow, the heterogeneous nature of soils, and crop response to the level of irrigation. Brown et al. used simulated annealing for on-farm irrigation scheduling considering seasonal water limits [77]. The objective was to maximize farm profit and was evaluated with a time-series simulation based on realistic plant growth models. Smout and Gorantiwar presented a water allocation linear programming model for optimizing the use of irrigation water to a medium irrigation scheme in India [84]. The model captured the deficit irrigation for each crop-soil-region combination. Yamout and El-Fadel developed a linear programming for setting policies for optimal water resources allocation on a regional scale [85]. Based on their study, the factors that greatly affect the water allocation scheme include profitability, public acceptability, and the effect of resources depletion.

It should be noted that most of existing studies focus on large scale management, such as, optimal irrigation and crop management on regional scale, and optimal scheduling for irrigation reservoir system. However, optimal on-farm level planning and irrigation scheduling remain a challenge from the research and practical perspectives [3]. For the studies focused on-farm level management, the granularity is typically a whole farm level, such as irrigation scheduling and crop rotation for the entire piece of land. Additional investigations are necessary to study the effect of the precision levels for on-farm management. In summary, majority of the literature focus on maximizing economic benefit, while maximizing yield and water use effectiveness were also adopted in several studies. Crop selection and irrigation management are among the main decisions to be made. Realistic constraints such as seasonal water limits, the heterogeneous nature of soils, and crop response to the level of irrigation applied are often considered. In this study, the proposed model aims at maximizing economic benefit by applying optimal decisions on crop selection and irrigation management. Seasonal water limits, soils features are considered. In addition, spatial structure and management scales are also considered in the proposed model to achieve a farm-level precision land management.

Corn, which is widely used for grain processing, food, beverages, livestock feed, and ethanol, takes up to one-third of cropland in the U.S. and is the nation's biggest crop economically. Corn receives the most irrigation water overall of American crops:

approximately 19 billion cubic meters annually [86]. Eighty-seven percent of irrigated corn in the U.S. is grown in high or extremely high water stress regions such as the Great Plains and the Central Valley in California, and over half of it depends on groundwater from the over-exploited High Plains aquifer. Extreme weather events due to climate change affect the corn industry significantly. For instance, irrigation water costs have soared to $0.89/cubic meter in 2015 from approximately $0.11/cubic meter in 2014 in the Fresno-based Westlands Water District due to severe drought in California. The devastating Midwest drought of 2012 drove corn prices to a record of $315/metric ton. These facts provided motivations for this study.

Motivated by the gap between theoretical decision making challenges and the pressing application need in reality, the objective of this study is to develop a mixed integer linear programming model to provide decision support for customized precision farmland management. In the proposed model, decisions for pre-season seed selection and irrigation scheduling are made based on management properties such as types of soil, spatial structure, and management scales under a series of realistic constrains. Careful consideration was given to the model framework so that it could easily account for weather stochasticity in the future.

The remainder of the paper is organized as follows: in Section 2, the problem statement for the farm-level precision land management model is presented. The basic mixed integer linear programming model is introduced in Section 3. The authors illustrate the method with a case study in California and discuss the extension and modification of the basic model in Section 4. Finally, the paper concludes with a summary of research findings and potential research directions in Section 5.

4.2 Problem Statement

Farmland management involves a sequence of planning and decision-making processes, the primary decision includes the scales and options of management. This paper focuses on solving two problems for farm-level precision land management. The first problem is to select the optimal crop management options within a customized management scale. The management options include seed type selection and irrigation frequency. The second problem is to choose the suitable management scale (size and shape) for these options. In other words, the model aims to assist the farmers to find the balance between precision level and management effort.

The "land unit" is defined as the minimum size over which management options are applied. The shape of a land unit is assumed to be square and the size of a land unit is informed by the measurement accuracy of soil types, agricultural working space, irrigation scale, and other physical limitations. Land unit could be viewed as the most precise block for a decision making level in farmland management. On the other hand, "decision unit" is defined as the farmer chosen scale for practical land management, which is a trade-off between convenience and precision. The size of decision unit could be any integer multiple of a land unit while the shape of a decision unit is a rectangle. A decision unit could be as small as one land unit or as big as the whole farmland. All the treatments and management options such as seed type selection and/or irrigation frequency setting in a decision unit are the same (among all land units in that decision unit). Based on these definitions, the problems could be a restatement of how to choose the scale of the decision unit and how to make optimal management option decisions within each decision unit. The hierarchical structure between land units and decision units make the proposed model flexible such that it can be extended to a farm that contains multiple disjoint pieces of land as well as to apply it to larger scales.

Several assumptions were made in the proposed model. It is assumed that the irrigate system already exists and it could apply different management options for each decision unit. It is also assumed that soil types will only affect the ability of holding water; they have the same nutrition levels [87]. The amount of water used in each irrigation is based on soil types, and the soils will achieve their saturated level after each irrigation [88]. It is assumed that irrigation will stop when the crops are dead. It should be noted that additional spatial constraints are included in the case study section to achieve a comprehensive analysis.

4.3 Model Formulation

In this section, the mixed integer linear programming model for farmland management problem is introduced. The objective is to maximize the farmer's annual net profit when considering a specific farm.

4.3.1 Mathematical notations

The mathematical notations are summarized in Table 4.1.

Table 4.1 Notations for proposed model

Subscripts		
r	$1,2,\dots,R$	irrigation frequency
s	$1,2,\dots,S$	seed type
$i(r,s)$	$1,2,\dots,I$	management option
j	$1,2,\dots,J$	land condition (soil types)
m	$1,2,\dots,M$	location of land unit in the horizontal axis
n	$1,2,\dots,N$	location of land unit in the vertical axis
$u(m,n)$	$1,2,\dots,U$	land unit (and its location)
v	$1,2,\dots,V$	decision unit
Binary decision Variables		
x_{iu}	whether management option i is used in land unit u	
y_{ru}	whether irrigation frequency option r is used in land unit u	
z_{su}	whether seed type s is used in land unit u	
Parameters		
A	size of total farmland	
B_v	set of land unit in decision unit v	
E	size of land unit	
C^o	overhead cost (cash and non-cash)	
C^w	unit cost for water	
C^f_{ij}	fixed cost of each irrigation for management option i used for land condition j	
C^m_{ij}	other farm operating cost for management option i used for land condition j	
L_{ju}	land conditions j for land unit u	

Table 4.1 continued

W_{ij}	amount of water needed for irrigation when management option i used for land condition j
Y_{ij}	unit maize yield when management option i used for land condition j
Y	minimum yield requirement for the farmland
B^m	budget limit for other farming cost
B^w	budget limit for irrigation
R^b	unit revenue for selling biomass
P	unit market corn price
W^l	irrigation water limitation per season
W^p	unit pre-irrigation water amount
Z	objective value
α	residues index
β	sustainability factor
γ	water use efficiency

4.3.2 Objective function

The objective is to maximize the farmer's annual net profit, which is defined as the total revenues subtracted by total system costs during the farming process. The binary decision variable x_{iu} represents whether management option i is used in land unit u. The total revenues include revenue from selling crop grain as well as net revenue from selling the by-product crop residues. For example, corn stover, which is the residue after harvesting the corn grain, is an important feedstock for production of second generation biofuels [18]. Alpha (α) is the residues index that is defined as the mass ratio between crop grain and biomass residues. Beta (β) is the sustainability factor which is the percentage of biomass residue that has to be left in the field to sustain the soil nutrients. Evapotranspiration, also known as crop water use, is defined as the water removed from the soil by evaporation from the soil surface and transpiration by the plants. Evaporation can account for 20% to 30% of growing season evapotranspiration. Gamma (γ) is defined as the water use efficiency, the ratio between evapotranspiration and total purchased irrigation water. Table 4.2 summarizes the mathematical formulation of components in the objective function.

Table 4.2 Components in the objective function

Component	Mathematical formulation
Crop sales revenue	$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}Y_{ij}EP$
Residue sales revenue	$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}Y_{ij}E\alpha(1-\beta)R^{b}/(1-\alpha)$
Other farming operating cost	$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, C_{ij}^{m}L_{ju}E$
Water purchasing cost	$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}W_{ij}C^{w}/\gamma$
Irrigation labor and equipment cost	$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}C_{ij}^{f}E$

A variety of system costs have been considered in the model including labor costs, irrigation costs, machinery costs, seed costs, chemicals costs, cash overhead, and non-cash overhead. Cash overhead consists of various cash expenses during the year that are assigned to the whole farm such as insurance, office expenses, machinery maintenance, and field supervisors' salary. Non-cash overhead includes capital recovery cost (annual depreciation and interest costs) for equipment and other farm investments. In order to have a concise expression and focus on the impact of irrigation water management, several costs including labor costs, machinery costs, seed costs, and chemicals costs are lumped into a single cost called "other farm operating costs". Irrigation cost includes water purchasing cost and a fixed cost of labor and equipment. C^{o} represents the overhead cost per acre (cash and non-cash). The objective function is thus defined as follows.

$$\max_{x_{iu}} Z = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}\big(Y_{ij}EP + Y_{ij}E\alpha(1-\beta)R^{b}/(1-\alpha)\big) - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}\big(C_{ij}^{m}E + C_{ij}^{f}E + W_{ij}C^{w}/\gamma\big) - C^{o}A$$

4.3.3 Constraints

The farming process requires upfront investment, which affects a farmer's cash flow. Farmers set up budget limits for certain cost categories. Constraint (1) ensures the total irrigation cost is below its budget. Constraint (2) ensures that other farm costs are below budget limit. No budget limit is set for overhead cost since it is independent from management decisions. For the consideration of food safety and a stable market, the government will encourage farmers to produce at least certain amount of crop in some cases. Similar total yield constraints are needed when there is a contract for a yield mandate. These situations are indicated in Constraint (3). Meanwhile, as a vulnerable and valuable resource, the amount of irrigation water is often limited in a growing season. This irrigation water limitation is reflected in Constraint (4). Constraint (5) ensures that the management decisions of land units are the same within a certain decision unit. Constraint (6) ensures that the irrigation frequency decisions are uniform within a certain decision unit. Constraint (7) ensures that the seed type selection decisions are uniform within a certain decision unit. It is noteworthy that the decision unit for irrigation is not necessarily the same as the decision unit for seed type. Only one type of decision could be made for each land unit, as indicated in Constraints (8) and (9). Constraints (10) and (11) govern that comprehensive decisions should be chosen from the union feasible region for each individual decision. The binary nature of decision variables are defined in the Constraint (12).

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}\left(C_{ij}^{f}E + W_{ij}C^{w}/\gamma\right) \leq B^{w} \tag{1}$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, C_{ij}^{m} L_{ju}E \leq B^{m} \tag{2}$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}Y_{ij}E \geq Y \tag{3}$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{u=1}^{U} x_{iu}\, L_{ju}W_{ij}/\gamma + W^{p}A \leq W^{l} \tag{4}$$

$$x_{iu} = x_{iu'} \quad \forall\, u, u' \in B_v \tag{5}$$

$$y_{ru} = y_{ru'} \quad \forall\, u, u' \in B_v \tag{6}$$

$$z_{su} = z_{su'} \quad \forall\, u, u' \in B_v \tag{7}$$

$$\sum_{r=1}^{R} y_{ru} = 1 \quad \forall u \tag{8}$$

$$\sum_{s=1}^{S} z_{su} = 1 \quad \forall u \tag{9}$$

$$x_{iu} \leq y_{ru} \quad \forall u \tag{10}$$

$$x_{iu} \leq z_{su} \quad \forall u \tag{11}$$

$$x_{iu},\ y_{ru},\ z_{su} \in \{0,1\}\ \forall i, r, s, u \tag{12}$$

4.4 Case Study

In California, the total area planted for field corn was 210 436 hectares (520,000 acres) with the highest corn grain production occurring in Central Valley. Meanwhile, the overextended Central Valley aquifer is one of the most vulnerable water resources, which could create additional risks for the $65 billion-a-year corn industry [86]. As an irrigated summer crop, the amount of irrigation applied to California field corn will largely determine how much water is available to the crop. Thus, it is imperative to implement precision farm management in this area. In this section, a farm located in Yolo County, Central Valley, California, is selected to conduct a case study and illustrate the proposed model. The size of the land is 65.56 hectares (162 acres) and the shape of the land is square. Extensions of the basic model on different implementation conditions are also discussed.

4.4.1 Data source

In the Central Valley, corn planting occurs from March through June and the time to mature is about 80 days to 130 days depending on the variety. Broadly, corn development can be divided into the vegetative stage that lasts through tassel and the reproduction stages that include silking, pollination, and grain filling. Since the plants don't consume much water in the early vegetative stage (first 4weeks) and do not need much irrigation, this study only considers the reproduction stages which involves irrigations (approximately 15 weeks). A variety of soil textures are present in the farms used for field corn production. Sandy soils are preferred for early plantings because they warm rapidly in the spring. Heavier soils are productive, provided they are well drained and properly irrigated. The soil information up to 0.91 meters (36 inches) in depth is collected using the Web of Soil Survey. This information is used to define six integrated soil types (sand, loamy sand, sandy loam, loam, silt loam or clay loam, and clay) based on the Unified Soil

Classification System (USCG). The water holding capacity of the soil types are adapted from literature [89, 90]. As shown in the upper part of Figure 4.1, there are five different types of soil in this farmland. Type 1 is the sand soil and Type 5 is the clay soil. This piece of land is divided into 324 land units and each land unit is a square with an acreage of 0.2 hectares (0.5 acres). If there are more than one soil type in a land unit, majority vote is applied to decide the soil type for that land unit, as shown in the lower part of Figure 4.1.

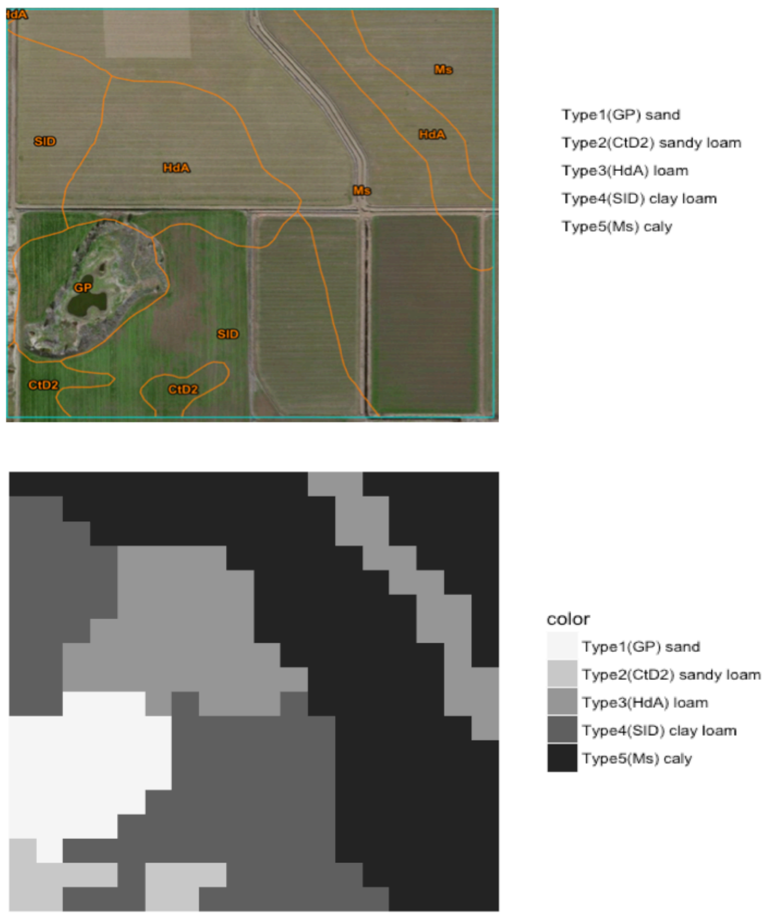


Figure 4.1 Satellite map (upper) and integrated map (lower) for soil types

The Central California Irrigation District (CCID) is one of the largest irrigation districts in the Central Valley, serving over 1,600 farms across more than 57870 hectares of prime farmland. The price of irrigation water is volatile and varies significantly by location, water usage, and water type (well water or surface water). In this study, farmers use both well and surface water at an average price of \$0.073/cubic meters (\$90/acre-foot). Seasonal irrigation water limits are set when insufficient water is available due to weather conditions and government regulations. For example, the CCID set seasonal irrigation water limits to be 2664 cubic meters (2.16 acre-feet) in 2014 and 3700 cubic meters (3 acre-feet) in 2015. The baseline of total water available is set to be 3083 cubic meters (2.5 acre-feet) per season in the case study. Six irrigation frequencies are available for selection (every day, every week, every other week, every three weeks, every four weeks, and never). Irrigation cost and overhead cost information are based on estimates from the Natural Resources Conservation Service (NRCS) and University of California Cooperative Extension [91]. Currently, almost all corn grown in California is irrigated by surface irrigation. In this study, the surge irrigation system is used with a water use efficiency of 0.6, meaning that 40% of the purchased water is lost during transportation, irrigation, and soil penetration. A pre-irrigation of 822 cubic meters (8 acre-inches) is applied in March. Other farm operating costs are estimated as \$1333/hectare (\$358 for machinery, \$91 for labor, and \$884 for Seed, chemicals); and these costs are uniformly applied [92].

Researchers from the University of California, Davis, reported a yield range from 12.54 to 18.81 metric ton per hectare with a minimum 1131 cubic meters survival water requirement for corn. In this analysis, twelve candidate grain corn seeds are created: three seeds for each of four major seeds types, including stringy, drought, smart, and extravagant.

These seeds have different levels of drought resistance and have a yield range from 13.17 to 18.81 metric ton per hectare [93]. These seeds share the same time needed to mature with a total evapotranspiration of approximately 63.5cm. The planting density is on average 83950 per hectare for each seed type. The average annual price received by U.S. corn producers from marketing years 2000 to 2015 is $141/metric ton, with a range from $71.65/metric ton to $271.26/metric ton according to the National Agricultural Statistical Service (NASS) of the U.S. Department of Agriculture. The baseline for corn market price in the case study is set at $141/metric ton. Corn stover could be used to serve as an abundant source of winter feed for cattle, and can also be used as the feedstock for biofuel production. The annual corn stover yield is estimated based on corn grain yield with a residue harvest index of 0.5, meaning 50% of the above ground biomass is grain and the amount of corn stover is the same as grain [68]. Papendick and Moldenhauer [94] showed that a 30\% removal rate results in 93\% soil cover after residue harvest. Thus, the sustainability factor (β) is set to be 0.3. It is assumed that the farm under consideration does not have a baler and therefore prefers to sell unharvested stover and let the buyer do harvesting. The unit revenue for selling unharvested cornstalks is $35 per metric ton [95]. All cost data have been adjusted for inflation to 2015 U.S. dollars.

4.4.2 Results for Model I

Model with Constraint (1) to Constraint (4), and Constraint (12) is defined as the Model I. The objective function and major constraints are consistent with the literature [7, 84]. In Model I, the size of a decision unit is set to be equal to the land units. Spatial structure and management scales are included in other models which will be defined later. Figure 4.2 shows that the managing option decisions are mainly chosen by the soil conditions; sandy land needs irrigation more frequently while clay land needs less irrigation. All decision unit chose the same seed type.

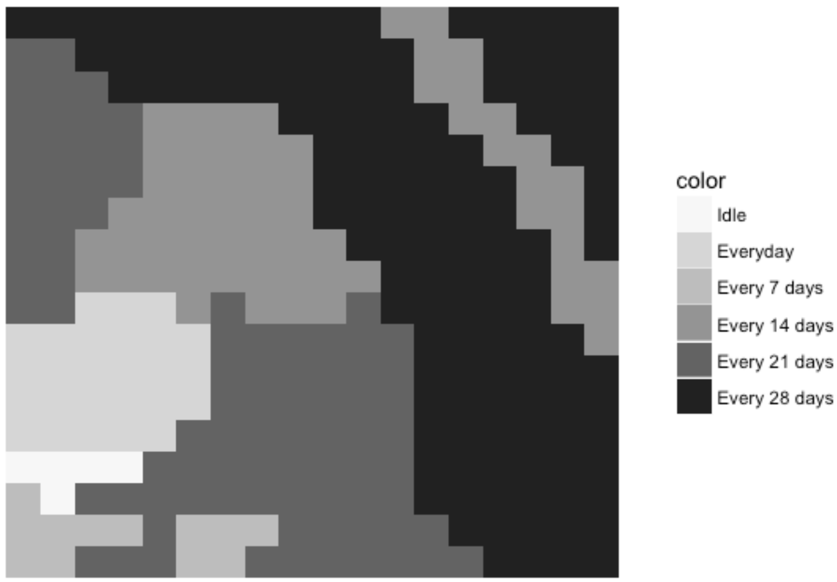


Figure 4.2 Irrigation decisions for basic model

These results are consistent with a "farming by soil" philosophy [3]. Part of the sandy land is idle due to the total irrigation water amount limitation. The net profit for this 65.56 hectare of farmland is $29,615, which yields to a marginal profit of $451.85/hectare. In order to have a baseline for comparing with previous literature and different model settings, a baseline scenario is introduced. In the baseline scenario, the size of a decision unit is set to be the entire land (uniform decisions for the whole farmland). The model yields to an average profit of $113.22/hectare under this scenario. University of California

Cooperative Extension reported an average profit range of $72.65/hectare to $135.91/hectare under same corn price with similar conditions [91]. The baseline scenario's profit located at higher part of this range. The average profit of $113.22/hectare from baseline scenario will be used for comparison between difference models.

Comparing results from Model I with the baseline scenario, although the Model I increase the profit significantly, these results require the most precise level of management, for example, valves in the surge irrigation system need to be switched at each irrigation. Model I should be regarded as the practice with highest precision requirements, which will serve as the upper bound on profitability.

4.4.3 Risk analysis

It should be noted that selection of modeling parameters is critical for the analysis results. In reality, the parameters in the model can exhibit great uncertainty due to market fluctuations, and extreme weather events. Sensitivity analyses, which consider the influence of one parameter on the objective at a time by assuming other parameters as constant, have been adopted as a paradigm to evaluate uncertainties in the parameters and their influence [11]. The parameters under investigation include corn market price P, irrigation water price C^w, other farm operating cost C_{ij}^m, overhead cost C^o, water use efficiency γ, and seasonal water limit W^l. The ranges of corn market price P and irrigation water price C^w are based on historical data, while the ranges of other parameters under investigation are $\pm25\%$ of the base level.

As shown in Figure 4.3, the parameters with largest impact on annual net profit are corn market price P and irrigation water price C^w. The significant variation of these two parameters leads to high leverage for the annual net profit. Corn market price is influential because it is the key factor for gross income. The trigger price of corn market price for growing is \$115.35/metric ton; corn market price lower than this point will lead to insufficient profit to cover farm costs. On the other hand, the termination price of irrigation water price is \$0.28/cubic meters; irrigation water price that is higher than this point will make farming unprofitable. Extremely high irrigation water prices due to special weather events will affect net profit significantly. The annual net profit is also sensitive to other farm operating cost C_{ij}^m, and water use efficiency γ, which gives us insight about potential directions to increase annual net profit.

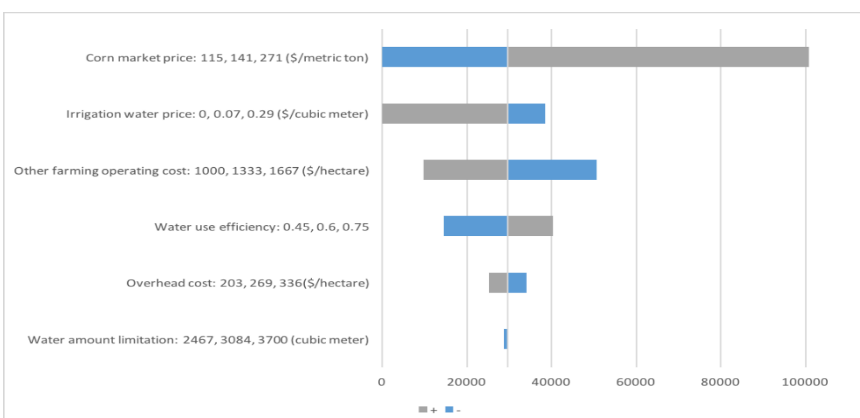


Figure 4.3 Sensitivity analysis of model parameters on annual net profit

Due to market fluctuations and climate change, making decisions under specific scenarios became a widely concerning problem. In 2014 and 2015, several California irrigation districts could not provide irrigation water for Class II lands, which refers to soils with moderate limitations that reduce the choice of plants or require moderate conservation practices. Farmers of these lands have to pay for private well water at an auction price over \$0.41/cubic meters, and the water suffers a loss factor related to the field's distance from the well source. A third of the Westlands district's farmland (242811 hectares) were left unplanted in 2014 due to especially high irrigation water prices. The local government asked the farmers to conduct risk analysis before making decisions [96]. A risk analysis tool based on our basic model could be easily applied to these farmlands and give appropriate recommendations. The analysis shown already indicates that corn market prices and irrigation water prices are dominating parameters for annual net profit. The simultaneous change of corn market prices and irrigation water prices by assuming other parameters hold constant can give us insight about the profit region. As shown in Figure 4.4, Region A is the nonprofitable region and Region B is the profitable region. If the speculated corn market price is relatively low and the irrigation water price is relatively high, farmers should change the crop type or leave the land idle to avoid further loss.

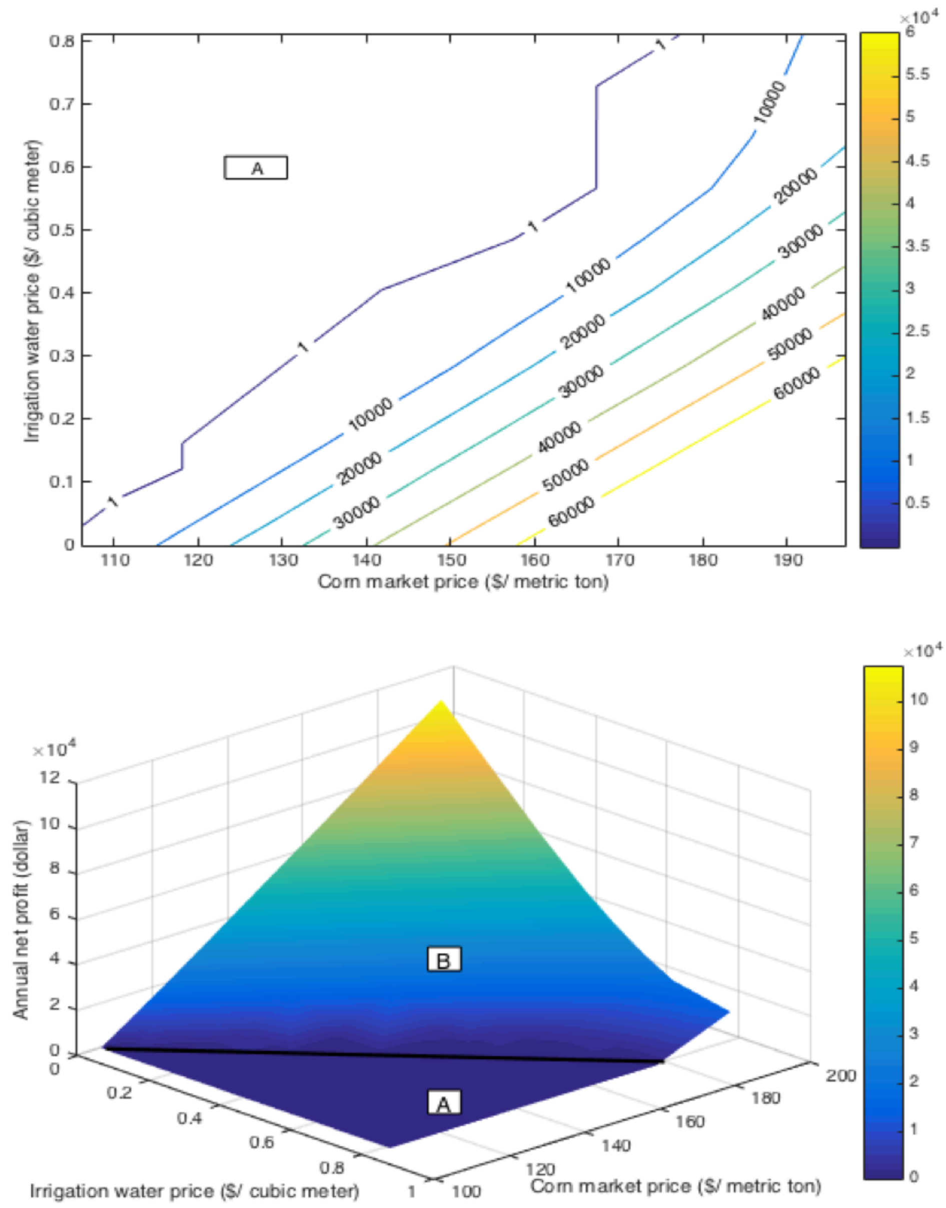


Figure 4.4 Contour plot (upper) and surface plot (lower) for profit region, Region A is the nonprofitable region and Region B is the profitable region, the darkness indicates the profit level

4.4.4 Discussion

In this section, some special cases are discussed to illustrate the flexibility of the basic model. Additional constraints are included to make the model robust under different realistic assumptions such as spatial structure and management scales.

Spatial relationship among land units are an essential part of the farmland model. In this section, the effect of size and shape of decision unit on annual profit are discussed. Even though precise farmland management will lead to more profit, it requires more management effort. The precision level and standardization level constitute a pair of tradeoffs. It is more realistic to make decisions on a larger scale and on a regular shape. In other words, each decision unit should contain multiple neighboring land units, and all land units in the same decision unit should share the same seed type and irrigation frequency. Three decision unit shape structures are investigated, namely square structure, row structure, and column structure. Meanwhile, several decision unit sizes are considered in order to find out the effect of scales.

Constraint (5) is added to the Model I; this new set of constraints ensure that the management decisions are uniform within a certain decision unit. This new constrained model will be referred to as Model II. Model I is a special case of Model II with the scale of decision units equal to the land unit.

Sixteen scenarios are generated based on the size, shape, and number of decision units. The first scenario, which applies uniform decision for the whole farmland, is the theoretical lower bound for this model and serves as a baseline for comparison. The gain ratio is defined as the annual net profit ratio between a certain scenario and the first scenario. Table 4.3 summarizes the annual net profit and gain ratio for each scenarios. As increasing the numbers of decision units (decreasing the size of decision units), the gain ratio is increased in general. These results indicate that detailed precision farmland management will bring high net profit. It also shows that the square structure is preferred because it has a higher gain ratio for the same size of decision unit and it has better flexibility. Table 4.3 also shows that the marginal benefit (of having more decision units) decreases. A highest gain ratio of 3.99 could be achieved by applying the philosophy of precision farm management.

Table 4.3 Effect of decision unit

Scenario	Number of decision unit	Shape of decision unit	Net profit (dollar)	Gain ratio
1	1	Square	7423	1.00
2	2	Row	9954	1.34
3	2	Column	8003	1.08
4	3	Row	11009	1.48
5	3	Column	15157	2.04
6	4	Square	20672	2.78
7	6	Row	13765	1.85
8	6	Column	15157	2.04
9	9	Row	16104	2.17
10	9	Column	16516	2.22
11	9	Square	20816	2.80
12	18	Row	15918	2.14
13	18	Column	16516	2.22
14	36	Square	27148	3.66
15	81	Square	27915	3.76
16	324	Square	29615	3.99

Up to this point, the model assumed that the decisions about seed type selection and irrigation frequency design are made simultaneously within each decision unit. However, due to the limitation of irrigation system, applying different irrigation frequencies to each decision unit may be cumbersome. Motivated by this practical limit, now the model allows different precision levels (size and shape of decision unit) between seed type selection and irrigation frequency design. Variables y_{ru} and z_{su} are introduced to indicate that each decision unit for seed type selection could have multiple irrigation frequencies and vice versa. Constraints (6) to (11) are added to Model II, and this new model is referred to as Model III. Model II can be viewed as a special case of Model III with a certain irrigation pattern.

Although Model III could capture any regular size and shape of decision unit in theory, three special irrigation patterns are investigated considering the irrigation system limitation, namely, Pattern 1: same irrigation frequency for each row (contains eighteen land units); Pattern 2: same irrigation frequency for each column (contains eighteen land units); and Pattern 3: same irrigation frequency for the whole farmland.

For each irrigation pattern, sixteen scenarios of seed type precision levels are investigated. These scenarios have the same definitions as in Model II. One dimension of precision management could be applied using Model III. On one hand, the authors want to find out the effect on annual profit by changing the precision level of seed type alone under certain irrigation patterns. On the other hand, the authors also want to investigate the effect on annual profit by changing the precision level of irrigation frequency alone under certain precision level of seed type selection. Under the same precision level of seed type selection, "relative gain for customized irrigation" (RGI) is defined as the net profit ratio between the highest profit (from Pattern 1, Pattern 2 or Model II) and profit for Pattern 3.

Table 4.4 summaries the annual net profit for three irrigation patterns under sixteen scenarios of seed type precision levels. As shown in the last column of Table 4.4, the RGI ranges between 2.00 to 2.42, which means that if the farmers have already decided the precision level of seed types selection, approximately 100% to 142% increase of net profit could be achieved by applying customized irrigation management.

Table 4.4 Effects of special irrigation patterns

Scenario	Net profit (dollar)			RGI
	Pattern 1	Pattern 2	Pattern 3	
1	15529	16516	7423	2.22
2	15041	16516	7423	2.22
3	14484	16516	7423	2.22
4	14814	16516	8100	2.04
5	15529	16516	7423	2.22
6	15529	16516	8781	2.35
7	15403	16516	8100	2.04
8	14590	16250	7423	2.19
9	16104	16516	8184	2.02
10	14610	15411	7592	2.18
11	15878	16533	9681	2.15
12	16104	16403	8201	2.00
13	15894	16516	7522	2.20
14	17345	15882	11235	2.42
15	17718	18150	12111	2.30
16	19001	18265	13404	2.21
Best gain ratio	2.56	2.46	1.81	

These increases are more significant under square decision units for seed type selection. On the other hand, if the farmers only allow precision management on seed type selection and use uniform irrigation frequency for the whole farmland (Pattern 3), the best gain ratio is 1.81. This result indicates that there is limited room for improving the net profit if farmers do not allow precision management on irrigation. When farmers allow some degree of precision management on irrigation (Pattern 1 and 2), the gain ratio will reach its upper bound at approximately 2.5.

To find out the quantitative relationships between the annual net profit with the number and shape of decision units under each irrigation pattern, regression analyses is conducted. Based on the hereinabove data analysis, logarithmic functions could be used to capture the effect of increasing the number of decision units, and a square structure has higher annual net profit under similar conditions. The following linear regression model is selected because it fits the data well and is easy to interpret.

$$P^a = \beta_0 + \beta_1 \ln(n) + \beta_2 \ln(n)\, I(Shape = square) + \varepsilon \qquad (12)$$

The response variable P^a is the annual net profit and the explanatory variable n is the number of decision units. ε is the random error that is not captured in the regression model, which is assumed to follow a normal distribution with mean zero and variance σ^2. $I(*)$ is the indicator function which takes value one when conditions are met and takes value zero when conditions are not met. β_0 could be interpreted as the baseline of annual net profit when there is only one decision unit. β_1 could be interpreted as the increment of annual net profit when the natural logarithm of the number of decision units increases by one. This increment will change to $\beta_1 + \beta_2$ when a square structure is selected. The best linear unbiased estimates and coefficient of determination(R^2) are summarized in Table

4.5. These results show that choosing a square structure and having more decision units has a positive effect on the annual net profit. The effects of number and shape of decision unit are more significant when two dimensional precision management is applied. A logarithmic function could describe the accelerated decline of the effects from the number of decision units on the annual net profit quite well.

Table 4.5 Summary of regression analysis

Parameters	Model II	Pattern 1	Pattern 2	Pattern 3
β_0	10245.2	14707.7	16319.62	7416.06
β_1	2233.4	344.3	65.59	184.06
β_2	1747.2	363.8	239.25	864.08
σ	2583	502.6	275.7	275.3
R^2	0.8738	0.8573	0.8131	0.9815

In summary, farmers could gain an additional 10%-80% net profit by employing precision management on seed type selection under certain irrigation patterns, and farmers could gain as much as an additional 142% net profit by working precision management on irrigation under certain seed type selection policy. One-dimensional precision management is relatively easier to implement but has a lower net profit. Precision management for irrigation appears to be more beneficial.

Besides confirming the dominant effect of crop prices and yields on net profit as stated in literature [91], this case study shows that irrigation water price, spatial structure, and management scales are also influential factors. The results from this case study show great economic potential of precision farmland management, and this recommendation is consistent with the literature [77].

- **Potential for sustainable water usage**

Although this study is mainly focused on economic analysis, it is important to take environmental issues into consideration. Water resources are limited and vulnerable, and corn is a thirsty plant. General strategies for coping with limited water include deficit irrigation of crops which can be stressed without significant loss of yield or quality, improving irrigation efficiency, improving crop genetics to develop varieties more tolerant to water stress, or planting other crops.

Reducing water amounts below what is required for corn will result in biomass reduction and grain yield reduction. What is more, the irrigation systems commonly used for corn in California do not allow close management of water stress. Thus, significant water savings can't be obtained by withholding water from the crop at present.

However, the method by which water is applied to the field could be improved. Strategies to maximize limited water include changes to irrigation management, design, or systems. Recall that in Model II and III, it is assumed that each land unit in a decision unit receives the same amount of water, which means some land units in a decision unit are over-irrigated and some water resources are wasted. Model I could eliminate this waste by having a land unit scale customized irrigation management.

Properly managed irrigation can apply a relatively uniform amount of water. However, application of high frequency may not be feasible with this system because of the labor input required for each irrigation. If farmers want to save water resources even further by applying deficit irrigation, new irrigation systems should be used such as sprinkler irrigation and traveling-gun irrigation. The proposed model could be easily modified to consider deficit irrigation. To illustrate this point, assume the irrigation technology could allow us to achieve at least partial saturation levels for a decision unit. Instead of assuming the corn cannot survive when it receives partial irrigation water, it is assumed that the yields of corn are depended by the saturation level. In other words, one more dimension of decision, the amount of irrigation water for each decision unit, are added in the model framework.

In summary, Model I is a special case of Model II with 324 decision units. Model II is a special case of Mode III with a certain irrigation pattern. These nested relationships indicate the flexibility of the proposed model.

4.5 Conclusions

In the study, a farm-level precision farmland management problem for pre-season seed type selection and irrigation water management is introduced. A mixed integer linear program is proposed with discussion on extensions and varieties of the basic model on different implementation conditions. Farmland in California serves as a case study to test the model's flexibility and economical optimality. The model gives qualitative descriptions and quantitative analysis for the management scale (number and shape of decision units). Special irrigation patterns are considered and the results show that the farmer's annual net profit could be significantly increased by applying one or two dimensional precision management decisions based on the proposed model. This model also serves as a decision making and risk analysis tool for farmers facing seasonal irrigation water limits and extreme drought conditions.

Note that this study is subject to a number of limitations. Firstly, the weather conditions such as temperature and rainfall are not considered in this model. These weather parameters affect the evaporation level of plants, pre-irrigation amount, and moisture level of the soil significantly. In addition, as discussed in the risk analysis, the parameters are not certain. Thus, a linear programming model with constant coefficients cannot fully describe the decision making environment [97]. Other modeling methods such as stochastic programming, dynamic programming, and robust optimization could be investigated [68]. In addition, multi-period models are needed for deficit irrigation design and invest new irrigation system.

The authors are working on a modified model which could take multi-period decisions of the seed hybrid and plant population selection, and amount of irrigation water, taking uncertain weather conditions and market price into consideration. This modified model would make the irrigation frequency and amount more flexible and precise. A stochastic program would be a natural fit to solve this problem; the first stage decision could be which type of plant seeds to grow while the other stage decisions could be the land management options such as irrigation amount for each irrigation.

In the case study presented to illustrate and validate this optimization model only considers a certain piece of land. However, the shapes of farmland could affect the agricultural machinery paths and the homogeneous features of the soil could affect the shapes and sizes of decision units [98]. Motivated by finite element analysis, other future work includes develop models that allow different shapes and sizes of decision units in a piece of land. Last but not least, the proposed model could be used to evaluate other crops as well; and the interaction among plants such as plant population, leaf cover, and water competition could be stressed in future research.

CHAPTER 5 A MULTI-STAGE STOCHASTIC PROGRAMMING MODEL FOR FARMLAND MANAGEMENT UNDER UNCERTAINTIES [4]

Abstract

Farmland management and irrigation scheduling are vital components of productive agricultural economy. A multi-stage stochastic programming model is proposed to maximize farmer's annual profit under uncertainties. The uncertainties under investigated include crop price, irrigation water availability, and precipitation amount. The first stage makes the pre-season decisions including the seed type selection and plant population selection, while the later stages determine when to irrigate and how much water should be used during each irrigation. The case study based on a farm in Nebraska show that a 10.22% profit increase could be achieved by taking corn price and irrigation water availability uncertainties into consideration using two-stage stochastic programming formulations. An additional 13.08% profit increase could be achieved by also taking precipitation amount uncertainties into consideration under multi-stage stochastic programming formulations. The stochastic model outperforms the deterministic model in the stochastic environment, especially when there is limited water supplies. These results

[4] This chapter of dissertation is preparing to submit to European Journal of Operational Research

indicate multi-stage stochastic programming is a promising way for farmland management under uncertainties and can increase farmers' income significantly.

5.1 Introduction

As the world population increases and area of arable land decreases, it becomes vital to improve the productivity of the available farmland. For thousands of years, drainage basins irrigation has been used to assist in the growing of agricultural crops, revegetation of disturbed soils in dry areas and during periods of inadequate rainfall. During recent decades, the advent of diesel and electric motors led to systems that could pump groundwater out of major aquifers and help increase the crops productivity. However, recent concerns have been raised regarding permanent loss of aquifer capacity, declining surface and groundwater supplies [99] and increased pumping costs [100]. Thus, decision making of management practices under limited water supplies is critical for sustainable agriculture and food security.

Corn is the most widely adopted row crop in the U.S and takes up to one-third of cropland nationwide. It has been mainly used as food, livestock feed, and bio-energy feedstock. Irrigated corn accounts for nearly 20% of total U.S. corn production while occupying only 15% of areas. Eighty-seven percent of irrigated corn in the U.S. is grown in high or extremely high water stress regions such as the Great Plains and the Central Valley in California, and over half of it depends on groundwater from the over-exploited

High Plains aquifer. Corn occupies more irrigated acres in these area than any other crops [100] and receives the most irrigation water among all of American crops [86].

Evapotranspiration (ET), also known as crop water use, is defined as the water removed from the soil by evaporation from the soil surface and transpiration by the plants. ET is driven by a tremendous drying force the atmosphere exerts on soil or plant surfaces. Hence the magnitude of daily ET will vary with atmospheric conditions. High solar radiation and air temperatures, low humidity, clear skies, and high wind increase ET, while cloudy, cool and calm days reduce ET. For example, reported seasonal corn ET averages around 24 inches in the humid eastern area of Nebraska compared to 28 inches for the more semi-arid southwestern region of the state [101]. ET is also affected by growth stage, length of growing season, soil fertility, water availability, and the interaction of these factors.

Deficit irrigation should be considerated where precipitation is low and irrigation water supply is restricted. Deficit irrigation refers to the method that distributes a limited amount of irrigation water to satisfy essential water needs of plants [79]. Reasons for limited water supplies include, but not limited to: restricted capacity of the irrigation well; restricted pumping allocations; reduced surface irrigation water supplies etc. When water supplies cannot fully compensate for crop ET, yields are reduced comparing to the fully irrigated crop. Under water-limited conditions, corn yields typically display a positive correlation with total seasonal water use. Grassini et al found that there is a linear relationship between potential grain yield and seasonal ET, and this relationship is valid across a wide range of grower fields and climatic conditions located in south-central

Nebraska [102]. On the other hand, applying additional irrigation beyond seasonal ET requirements can lead to leaching and/or water left in the soil. The impact of water stress on corn grain yields varies significantly with crop growth stage.

For corn, the growth stage is divided into five stages: establishment, vegetative, flowering, grain filling and ripening. Corn is relativity insensitive to water deficits during early vegetative growth and ripening periods because water demand is relatively low. Plants can adapt to water stress throughout most of the vegetative period to reduce its impact on grain yield [103]. However, corn is much more sensitive to water stress from flowering through grain filling stages [103-105]. Severe water deficits during the silking and pollination process of the flowering stage will cause silk drying, which will lead to little or no grain yield. In addition, insufficient water during the grain filling stage may result in reduced yield due to a reduction in grain size. On the other hand, waterlogging should be avoided, particularly during the flowering and grain filling stages.

Key factors that affect the irrigation management decisions include soil characteristics, plant features, irrigation methods, and atmospheric factors. Soil characteristics such as water holding capacity and infiltration rate could affect water movement and root penetration. In addition, some root-restricting layers at shallow depths can also restrict root development. Water consumption related plant phenotype includes features like crop development time, rooting depth, and seasonal crop water use. These features will affect the drought tolerence. Selecting the appropriate plant population is as important as choosing the suitable seed type. Lower population could reduce the transpiration component by the crop of ET and require less precipitation and irrigation. Irrigation methods determine irrigation water use efficiency. Center pivot sprinkler systems can achieve a efficiency of up to 90 percent. However, conventional gated pipe irrigation system has only a 50 percent water use efficiency, meaning that half of the water is lost during the irrigation process. As a nature source of water for farmland, when and how much will the precipitation occur is another key issue for irrigation scheduling. Moreover, factors like crops price, precipitation amount, and irrigation water availability are not deterministic in real world application. Farming activities are highly affected by these uncertainties. Thus, optimization tools for farmland management and irrigation scheduling are needed under uncertainties.

Mathematical programming has been widely used in farmland management, especially in irrigation management. Sabu et al used a multi-level approach based on dynamic programming to find optimal irrigation allocation on a reginal scale [9]. Brown et al used simulated annealing for on-farm irrigation scheduling considering seasonal water limits [10]. Georgiou and Papamichail used simulated annealing and a gradient descent algorithm for irrigation reservoir and crop planning optimization [5]. Their method accounted for variable reservoir inflows and climate variability for crop planning. Ganji et al proposed a constraint state formulation for stochastic control of the weekly deficit irrigation strategy [11]. The model is based on the first and second moment analysis of the stochastic soil moisture state variable and consider the crop water demands uncertainties. Although these studies contain some sort of uncertainties, farmland management and irrigation scheduling under uncertainties such as crops price, precipitation amount, and irrigation water availability have not been studied extensively.

Stochastic programming is a mathematical programming method where some of the parameters incorporated into the objective or constraints are uncertain. It could reflect the dynamic variations of system conditions, especially for sequential decision making problems. Stochastic programming has been adopted in water management on reservoir system for decision making under uncertainties. Pereira and Pinto proposed a stochastic programming framework to minimize the expected operation cost for interconnected reservoir system under uncertainty [6]. Huang and Loucks developed an inexact two-stage stochastic programming model for water resources decision making under uncertainty [7]. Li et al. extended this work to an inexact multi-stage stochastic programming model [8]. However, to the best of the authors' knowledge, few applications to farmland scale irrigation management based on stochastic programming were reported. Therefore, the feasibility and advantage of modeling farmland management problem via stochastic programming should be investgated, which also motivated this study.

In summary, the effect of limited water on corn grain yield is significant and appropriate decisions are needed to optimize farmers' profits, particularly under stochastic environment. Factors like weather conditions, market price, soil characteristics, plant features, and irrigation methods should be all taking into consideration when choosing irrigation and agronomic practices. In this study, a multi-stage stochastic programming model is formulated considering uncertainties such as crops market price, precipitation amount, and irrigation water availability. The first stage makes the pre-season decisions including the seed type selection and plant population selection, while the later stages determine the irrigation schedule.

The remainder of this chapter is organized as follows. The problem statement is presented in Section 5.2. In Section 5.3, the model formulations are introduced. A case study is conducted to illustrate and validate the optimization model in Section 5.4. Finally, we conclude the chapter in Section 5.5 with a summary and potential research directions.

5.2 Problem Statement

In this study, a multi-stage stochastic programming model is formulated considering uncertainties such as crops price, precipitation amount, and irrigation water availability. These uncertainties are represented by scenario trees as realization of probability distributions or stochastic processes. The objective is to maximize the farmer's annual net profit by finding the optimal decisions for seed selection and irrigation schedule. There are nine time period ($t = 0,1,\ldots,8;$) considered in the model. The time period 0 ($t = 0$) is at the beginning of the year, the time period 1 ($t = 1$) is at the beginning of the corn flowering stage. The time period 1 to 8 ($t = 1,\ldots,8;$) corresponds to the eight weeks for the flowering and grain filling stages of corn. Crop price and seasonal irrigation water availability information are assuming to be released at the beginning of time period 1 ($t =$

1). Precipitation information of these eight weeks are available at the end of 1st to 8th time period.

The decision maker has to take a sequence of decisions at each time period in order to maximize profit. In stochastic programming framework, the decision maker makes some decisions at the first stage. The outcome of these decision will be affected by some random events, the later stage recourse decisions could be made to adjust these effects. In other words, stochastic programming gives first-stage decisions and a collection of recourse decisions based on each random outcome. In this problem, decision maker makes the pre-season decisions including the corn seed type selection and plant population selection at the first stage ($t = 0$). At the beginning of second stage ($t = 1$), realization of corn market price and seasonal irrigation water availability become available and the second stage decisions of how much irrigation water should be put in the field for week one ($t = 1$) are made. At the beginning of 2nd to 8th time period, similar irrigation schedule decisions are made based on available information so far. The precisely decision process has the form:

decision of seed type and plant population→ observation of corn market price and seasonal irrigation water limits→ decision of irrigation for $t = 1$→ observation of precipitation for $t = 1$→ . . .→observation of precipitation for $t = 7$ →decision of irrigation for $t = 8$.

Crop yield response functions are employed to determine the deficit levels[106]. As shown in Equation 1, Y_m is the maximum crop yield under full irrigation, k_i is the crop yield response factor to water and is a function of the crop type and the stage of growth, I the total number of crop growth stages, ET_a the actual crop stage evapotranspiration, ET_c the crop stage evapotranspiration without water stress.

$$\frac{Y_a}{Y_m} = \prod_{i=1}^{I}[1 - k_i(1 - \frac{ET_a}{ET_c})_i] \quad (1)$$

For deficit irrigation of corn, it is suggested that water could be saved to the flowering and grain filling stages by reducing irrigation during the vegetative stage, since corn is much more sensitive to water stress from flowering through grain filling stage. It is assumed that the irrigation will only take place in flowering and grain filling stages. An integrated crop yield response factor for flowering and grain filling stage is used to make Equation 1 as linear function of irrigation. Decision maker could decide to apply less than normal water for each irrigation during the flowering and grain filling stages to maximize the farmer's annual net profit.

5.3 Model Formulation

The deterministic and stochastic models for this farmland management problem are introduced in this section. The deterministic model is firstly introduced as a baseline model and then the multi-stage stochastic programming model is presented to address decision making under uncertainties. The objective is to maximize the farmer's annual net profit.

5.3.1 Mathematical notations

The mathematical notations are summarized in Table 5.1.

Table 5.1 Mathematical notations

Subscripts		
r	$1,2,\dots,R$	Index for plant population levels
s	$1,2,\dots,S$	Index for seed type
$i(r,s)$	$1,2,\dots,I$	Index for pre-season management option
w	$1,2,\dots,W$	Index for scenario
l	$1,2,\dots,L$	Index for deficit levels
t	$1,2,\dots,T$	Index for time periods (during flowering and grain filling)
Decision Variables		
x_i	Whether pre-season management option i is applied, binary variables	
y_t	Irrigation water amount used during time period t, non-negative variable	
Dependent Variables		
z_t	Whether irrigation is given during time period t, binary variables	
d_l	Whether deficit level l in applied, binary variables	
Y_l^c	Actual yields under deficit level l, non-negative variables	
M_t	Water available in soil at the beginning time period t, non-negative variables, $M_1 = 0$	
ET_t^a	Actual evapotranspiration during time period t, non-negative variables	
L_t^W	Leaching water amount during time period t	
Parameters		
A	Total area of the farmland	
C^o	Overhead cost (cash and non-cash)	
C^{wt}	Unit cost for water	
C_i^s	Unit cost for seed under pre-season management option i	
C^f	Unit fixed cost of each irrigation	
C_i^m	Unit other farm operating cost for pre-season management option i	
D_l	Percentage of the maximum crop stage evapotranspiration achieved in deficit level l	
Y_i^m	Maximum unit crop yield when management option i used	
Y	Minimum yield requirement for the farmland	
B^m	Budget limit for other farming cost	

Table 5.1 continued

B^w	Budget limit for irrigation
G	Unit market corn price
P_w	Probability for each scenario
K_i	Crop yield response factor to water during flowering and grain filling for pre-season management option i
ET_t^m	The crop stage evapotranspiration without any water stress during time period t
R_t	Total precipitation during time period t
H	Soil water holding capacity
W^l	Total irrigation water limitation during flowering and grain filling season
W^p	Unit pre-irrigation water amount
M^b	A sufficiently large number used in big-M method
γ	Water use efficiency
ξ_t	A random vector and its particular realization at each time period

5.3.2 Deterministic model

A mixed integer linear programming model is formulated in this section. All the system parameters are assumed to be known with certainty in the deterministic model.

The objective is to maximize the farmer's annual net profit, which is defined as the total revenues subtracted by total system costs. The binary decision variables x_i represent whether pre-season management option i is used. The positive decision variables y_t represent how much irrigation water is used during time period t. The binary variables z_t, which are dependent on y_t represent whether irrigation is given during time period t.

A variety of system costs have been considered in the model including labor costs, irrigation costs, machinery costs, seed costs, chemicals costs, cash overhead, and non-cash overhead. Cash overhead consists of various cash expenses during the year that are assigned to the whole farm such as insurance, office expenses, machinery maintenance, and field supervisors' salary. Non-cash overhead includes capital recovery cost (annual depreciation and interest costs) for equipment and other farm investments. In order to have a concise expression and focus on the impact of irrigation management, several costs including labor costs, machinery costs, and chemicals costs are lumped into a single cost called "other farm operating costs". Irrigation cost includes water purchasing cost and a fixed cost of labor and equipment. C^o represents the overhead cost per acre (cash and non-cash). The objective function is thus defined as follows:

$$\max_{x_i, y_t}\{GA\sum_{l=1}^{L} Y_l^c - AC^{wt}\left(\sum_{t=1}^{T} y_t/\gamma + W^p\right) -$$

$$A\sum\nolimits_{i=1}^{I} x_i\,(C_i^s + C_i^m) - AC^f\sum_{t=1}^{T} z_t - C^o A\}$$

Constraint (a1) and Constraint (a2) are the period soil moisture continuity equations. For each time period, irrigation and precipitation will replenish the soil moisture while ET and leaching will consume water. Irrigation and precipitation plus current soil moisture should be less than soil water holding capacity and the extra water will leach and waste, this requirement is reflected in Constraint (a3).

$$M_t + y_t + R_t - ET_t^a - L_t^W = M_{t+1} \quad for\ t = 1,2,\ldots 7 \qquad \text{(a1)}$$

$$M_t + y_t + R_t - ET_t^a - L_t^W \geq 0 \qquad for\ t = 8 \qquad \text{(a2)}$$

$$M_t + y_t + R_t - L_t^W \leq H \qquad \forall t \qquad \text{(a3)}$$

Constraint (a4) is the definition of deficit level. In order to have a smooth change among deficit levels, 101 equidistant levels from 0% to 100% are used. Constraint (a5) and Constraint (a6) are the crop yield response functions for water use based on Equation (1). Only one deficit level should be selected, this requirement is present in Constraint (a7) using binary variables d_l. Constraint (a5) to Constraint (a7) together is the so called "only one out of L constraints much hold" case. Constraint (a8) ensure only the selected deficit level will lead to meaningful actual crop yields. For the computation consideration, the M^b should be as small as possible and it is set to be equal to $\max_l Y_l^c$.

$$\sum_{t=1}^{T} ET_t^a / ET_t^m = \sum_{l=1}^{L} d_l \, D_l \tag{a4}$$

$$Y_l^c - \sum_{i=1}^{I} x_i Y_i{}^m \big(1 - K_i(1 - D_l)\big) \leq (1 - d_l) M^b \quad \forall l \tag{a5}$$

$$Y_l^c - \sum_{i=1}^{I} x_i Y_i{}^m \big(1 - K_i(1 - D_l)\big) \geq (d_l - 1) M^b \quad \forall l \tag{a6}$$

$$\sum_{l=1}^{L} d_l = 1 \tag{a7}$$

$$d_l M^b \geq Y_l^c \quad \forall l \tag{a8}$$

As a vulnerable and valuable resource, the amount of irrigation water is often limited in the key growing stages. This irrigation water limitation is reflected in Constraint (a9). For the consideration of food safety and a stable market, the government will encourage farmers to produce at least certain amount of crop in some cases. Similar total yield constraints are needed when there is a contract for a yield mandate. These situations are indicated in Constraint (a10).

$$A \sum_{t=1}^{T} y_t / \gamma \leq W^l \tag{a9}$$

$$A \sum_{l=1}^{L} Y_l^c \geq Y \tag{a10}$$

Total times of irrigation are needed to calculate the fixed cost of labor and equipment for irrigation. These costs occur only if the irrigation water amount is above zero, as reflected in Constraint (a11). As shown in Constraint (a12), only one seed type and plant population could be selected. Constraint (a13) makes a conservative assumption that there is no water in soil at the beginning of first time period. Constraint (a14) controls the domain of variables.

$$M^b z_t - y_t \geq 0 \qquad \forall t \tag{a11}$$

$$\sum_{i=1}^{I} x_i = 1 \tag{a12}$$

$$M_t = 0 \qquad for\ t = 1 \tag{a13}$$

$$x_i, d_l \in \{0,1\}, y_t, Y_l^c, M_t, ET_t^a, L_t^W \geq 0 \qquad \forall i, \forall t, \forall l \tag{a14}$$

5.3.3 Multi-stage stochastic programming model

In this study, precipitation amount, irrigation water availability, and corn prices are selected as the stochastic parameters to be investigated. Scenario trees are used as an approximation of probability distributions or stochastic processes. Subscript w is used to represent index of scenario with corresponding probability P_w, and the subscript is also incorporated into the decision variables and parameters. The multi-stage stochastic programming model is formulated as follows:

$$\max_{x_i, y_{tw}} \left\{ -A \sum_{i=1}^{I} x_i \left(C_i^s + C_i^m\right) - C^o A + \sum_{w=1}^{W} P_w \left\{ AG_w \sum_{l=1}^{L} Y_{lw}^c - AC^{wt} \left(\sum_{t=1}^{T} y_{tw}/\gamma + W^p \right) - AC^f \sum_{t=1}^{T} z_{tw} \right\} \right\}$$

s.t.

$$M_{tw} + y_{tw} + R_{tw} - ET_{tw}^a - L_{tw}^W = M_{t+1,w} \quad \forall w, t = 1,2, \dots 7 \tag{b1}$$

$$M_{tw} + y_{tw} + R_{tw} - ET_{tw}^a - L_{tw}^W \geq 0 \quad \forall w, t = 8 \tag{b2}$$

$$M_{tw} + y_{tw} + R_{tw} - L_{tw}^W \leq H \quad \forall t, \forall w \tag{b3}$$

$$\textstyle\sum_{t=1}^{T} ET_{tw}^a / ET_t^m = \sum_{l=1}^{L} d_{lw} D_l \quad \forall w \tag{b4}$$

$$Y_{lw}^c - \textstyle\sum_{i=1}^{I} x_i Y_i^m \left(1 - K_i(1 - D_l)\right) \leq (1 - d_{lw}) M^b \quad \forall l, \forall w \tag{b5}$$

$$Y_{lw}^c - \textstyle\sum_{i=1}^{I} x_i Y_i^m \left(1 - K_i(1 - D_l)\right) \geq (d_{lw} - 1) M^b \quad \forall l, \forall w \tag{b6}$$

$$\textstyle\sum_{l=1}^{L} d_{lw} = 1 \quad \forall w \tag{b7}$$

$$d_{lw} M^b \geq Y_{lw}^c \quad \forall l, \forall w \tag{b8}$$

$$A \textstyle\sum_{t=1}^{T} y_{tw} / \gamma \leq W_w^l \quad \forall w \tag{b9}$$

$$A \textstyle\sum_{l=1}^{L} Y_{lw}^c \geq Y \quad \forall w \tag{b10}$$

$$M^b z_{tw} - y_{tw} \geq 0 \quad \forall t, \forall w \tag{b11}$$

$$\textstyle\sum_{i=1}^{I} x_i = 1 \tag{b12}$$

$M_{tw} = 0,\ t = 1 \qquad \forall w$ (b13)

$x_i, d_{lw} \in \{0,1\}, y_{tw}, Y^c_{lw}, M_{tw}, ET^a_{tw}, L^W_{tw} \geq 0 \qquad \forall i, \forall t, \forall l, \forall w$ (b14)

$y_{tw} = y_{tw'}, \forall w, w'$ for which $\xi^w_{[t]} = \xi^{w'}_{[t]} \qquad \forall t$ (b15)

The first stage decisions involve decisions which must be made before the uncertainties are realized. After the uncertainties are progressively realized, the later stage decisions are made. In this model, the first stage decision variables are x_i. The later stage decision variables are y_{tw}. Constraints (b12) is the first stage constraints; this constraint remain the same in all scenarios, and they are the same as in the deterministic linear program model. The rest of the constraints change based on the stochastic scenarios. We use the notation ξ_t $(t = 1, \dots, T-1;)$ to denote a random vector and its particular realization at each time period. The decision at each period ($t = 1, \dots, T$) depends from the realization of ξ_t up to time t. Generally, at stage $t = 1, \dots, T$, the scenarios that have the same history $\xi_{[t]}$ cannot be distinguished, so we need to enforce the nonanticipativity constraints by adding Constraint (b15).

5.4 Case Study

We apply the farmland management framework based on stochastic programming for a case study on a farm in Cherry County, Nebraska to illustrate and validate the optimization model. Half of harvested row crop production in Nebraska are irrigated (About 8 million acres), where corn occupies approximately 70 percent of the irrigated acreage [101]. Consequently, improving farmland management and irrigation scheduling have significant impact on the water resources and farmers' income.

5.4.1 Data sources

Conservative irrigation management typically assumes a three-foot effective root zone for field corn. The soil information up to three feet in depth is collected using the Web of Soil Survey. This information is used to define integrated soil types (fine sand, loamy sand, sandy loam, fine sandy loam, loam, clay loam, and clay) and water holding capacity of these soil types [107]. A farm of size 150 acreage in Cherry County, Nebraska is selected for analysis. As shown in Figure 5.1, 95% of the soil is loamy sand and 5% of the soil is sandy loam, both of them are coarse soil. The soil water holding capacity is assume to be 1.1 inch per foot for the whole land [107]. The irrigation water is supplied by center pivot sprinkler systems of 800 gallons per minute. The water use efficiencies for center pivots outfitted with low pressure drop nozzles are typically rated at 85% [108], meaning that 15% of the water is lost during transportation, irrigation, and soil penetration.

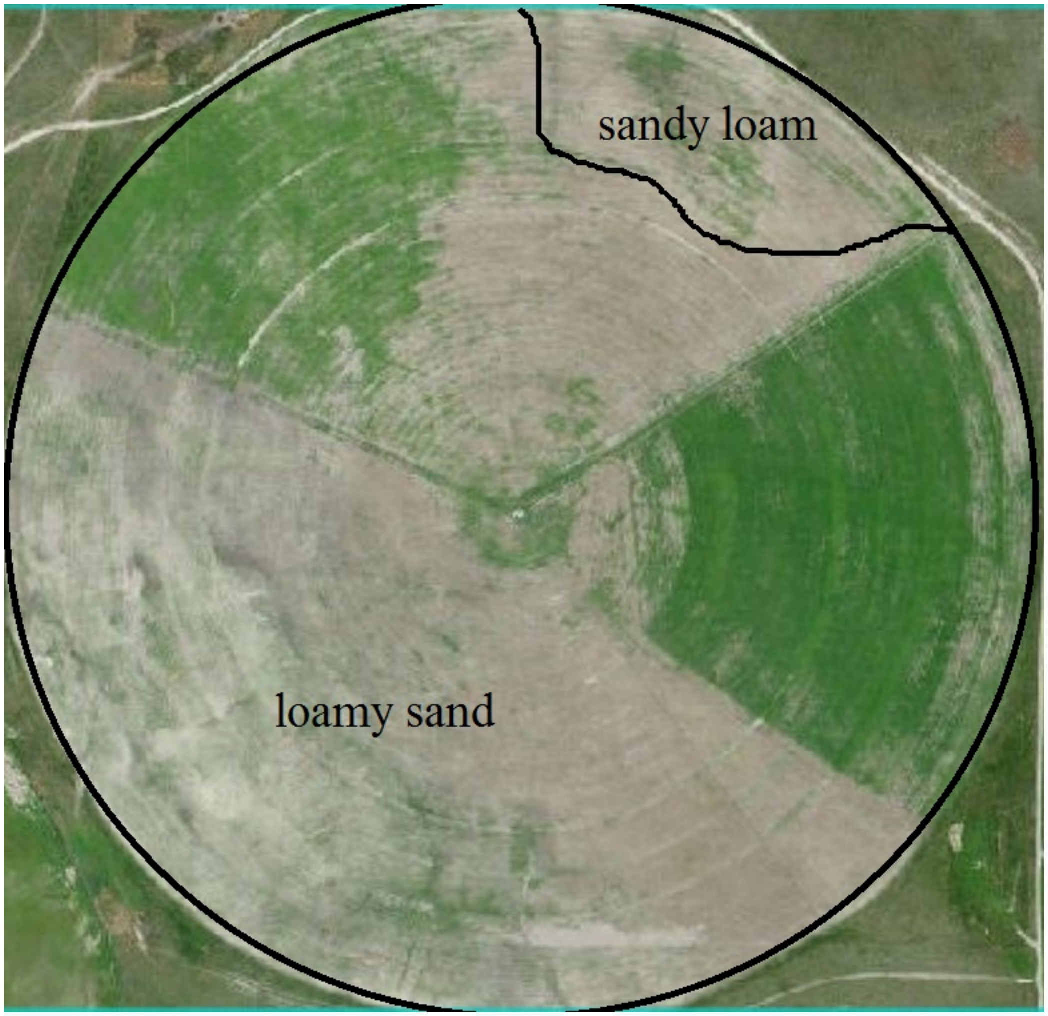


Figure 5.1 Satellite map of the selected farm

The root zone should be wetted at sowing in order to obtain a good germination rate and rapid root development. Thus, pre-irrigation at spring are needed to refill the soil profile, particularly when there is limited winter precipitation. Since corn does not consume much water in the vegetative stage and do not need much irrigation, this study focuses on the irrigation for flowering and grain filling stages (approximately eight weeks). The average crop water use for these period is range from 0.2 to 0.32 inches per day [101].

The price of irrigation water is volatile and varies significantly by location, water usage, and water type (well water or surface water). In this study, it is assumed that farmers use well water at an average price of $12/acre-inch. The other farm operating costs and fixed irrigation cost are adopted from the Nebraska Water Optimizer Single-Field Version (NWO) [109]. The seed features such as drought tolerance, target yields, and suggested plant population are based on commercialized crop hybrids. The maximum yields of these seeds under full irrigation range from 160 to 230 bushels per acre.

The corn prices received by U.S. corn producers from 2000 to 2015 were collected based on the National Agricultural Statistical Service (NASS) of the U.S. Department of Agriculture. The baseline for corn market price in the deterministic case is set at $3.6 dollar per bushels. Historical precipitation information of Cherry County is obtained from National Oceanic and Atmospheric Administration (NOAA)'s National Centers for Environmental Information (NCEI). Detailed discussions on the distribution of corn price, precipitation amount, and total water limits are given in the following scenario generation section. All cost data have been adjusted for inflation to 2015 U.S. dollars.

5.4.2 Scenario generation

Computational methods for solving stochastic optimization problems require a discretization of the underlying probability distribution of the uncertain parameters. In stochastic programming, scenarios describe possible values that the uncertain parameters may take. The scenario trees are approximations of (continuous) distribution functions because they contain only a limited number of outcomes. When the continuous probability distribution are used to represent the uncertainties, the Sample Average Approximation (SAA) method is often used to generate scenario, for instance via Monte Carlo sampling [110, 111]. When the historial data are available but it is not easy to fit them in some well-known probability distribution, moment matching method is often employed for scenario generation [35, 112].

There is two main requirements for choosing the size of scenario, the first one is that the number of scenario should be large enough so that they could represent the probability distribution; the second one is that the number of scenario should be relatively modest so that it could be solved with reasonable computational effort. Note that stochastic programming in general and multistage programs in particular have been known to be computationally challenging to solve. Enlarging the size of the scenario will generally achieve a better approximation. However, the size of the scenario tree directly impacts the computational complexity of stochastic programming models. The stability tests are used to test the stability of a scenario generation process for a given size of the scenario. For two-stage stochastic programming, the in-sample stability test is used to as a test of the internal consistency of a model (scenario generation process). The standard in-sample stability test and out-of-sample stability test is not suitable for multi-period trees, as the nodes beyond the root do not coincide [113]. The weak out-of-sample stability test for multiperiod trees is used to evalue the stability of scenario generation process. The procedure is building two scenario trees and find the corresponding solutions. Then solve the optimization model on the first scenario tree with the first stage decisions from the second tree, and vice versa. We should get approximately the same optimal objective values if the method is out-of-sample stable. In this case study, the size of scenario is set to be 200 given the computational power, stability test will be presented in the results analysis section.

Finding the correct distribution is also critical for scenario generation. Since we are considering a single year problem, a meaningful corn price should be the average price received by farmers after the corn is harvested and ready to sell. The market year of corn sales start at September, and six month sales season are considered. In other word, we want to find the distribution of average corn price from September to the next February. In order to take the most advantage of the available information, this distribution should be conditional on the corn price before the sowing season, which is April in Nebraska. Shapiro-Wilk normality test of the historical corn price data gives a P-value of 0.838, meaning that these conditional data follow normal distribution. The maximum likelihood method is used to get the parameter estimations. The mean is the corn price at April minus 0.147, and the standard deviation is 0.585. the distribution could be presented as $N(G_{Apr} - 0.147, 0.585)$.

As one of the most important weather variable, the methodology for precipitation predition is fairly well established and reliable simulation techniques are available [114, 115]. In this study, a two step process is adopted for precipitation generation: the daily precipitation occurrence (i.e. wet or dry day) is modeld upon a first order two state Markov chain and once it rains, the precipitation amount is assumed to follow gamma distribution [79, 116, 117]. It is assumed that each week's precipitation follows its unique gamma distribution and the simulation results are then sum up to weekly basis.

Based on a center pivot sprinkler systems of 800 gallons per minute capacity, the theoretically upper bound for eight weeks' total water availability is 2355 acre-inch. However, high application rates of water to coarse textured soils can destroy surface soil structure and enhance runoff. Thus, the practical upper bound for total water available is 2240 acre-inch [101]. The system down time due to maintenance, system failure, insufficient groundwater, and electrical load control should also be taken into consideration. For example, Nebraska Public Power Districts can be authorized to interrupt power up to six 12-hour periods during a week in the "anytime control" mode [107]. The lower bound is set to be 1649 acre-inch, or 70% of the theoretical upper bound. Since there is not much data to fit a distribution of total water limits, an uninformative uniform distribution with range 1649 acre-inch to 2240 acre-inch is assumed in this study.

Since the distributions of random variables are available, a common approach to generate the scenario to a manageable size is by using SAA method based on Monte Carlo simulation. It is assumed that these three random variables are independent. The individual scenarios in form of a fan are used as input for scenario tree construction and reduction based on Heitsch and Römisch's method [118]. The General Algebraic Modeling System

(GAMS)/SCENRED2 is utilized for scenario reduction and later solving the mathematical model.

5.4.3 Measures of information

Generally, stochastic programming takes advantage of taking more information about the future uncertainties into consideration when making decisions. Thus, measures of information are needed to discuss the value of stochastic programming and information. In two-stage stochastic programming, several approaches based on different levels of available information have been widely used in literature. The expected value problem solution (*EV*) is obtained by replacing all random variables by their expected values and solving a deterministic program. The expection of expected value problem solution (*EEV*), denotes the expected result of using the solution from the deterministic model EV to the stochastic environments. The wait-and-see solution value (*WS*), denotes the expected value of using the optimal solution for each scenario. The solution value of the stochastic model, also known as the here-and-now solution, denotes the optimal solution value to the recourse problem (*RP*).

For the maximization models in particular, the following inequalities are satisfied [119]:

$$EEV \ \leq RP \ \leq WS$$

There are two concepts mainly used for measuring the information for two-stage stochastic programming, namely, the expected value of perfect information (*EVPI*) and the value of the stochastic solution (*VSS*) [120]. In this context, the $EVPI = WS -$

RP compares here-and-now and wait-and-see approaches, a small $EVPI$ means a small addition profit when having prefect information. $VSS = RP - EEV$ compares the here-and-now and expected values approaches. A large VSS means that the approximation of using EV in the stochasitc environments in bad decisions.

The WS is still valid in multi-stage stochastic programming, where the decision maker assumes to knows at the first stage the realizations of all the random variables. However, the EEV for multi-stage stochastic programming is sometimes misleading. It can happen that the first stage solution in the EV problem performs better than the solution of the RP one, because the RP model contains nonanticipativity constraints in later stages, which are ignored (relaxed) when getting EEV [121]. One way to avoid this issue is using a chain of values $VSSt$ which takes into account the information until stage t of the associated deterministic model [121]. However, these results are valid if only the right hand side constraints are stochastic. Another way is to define the value of multi-stage stochastic programming (VMS) as the difference between the optimal objective values of the two-stage (v^{TS}) and multi-stage formulations (v^{MS}) [122]:

$$VMS = v^{MS} - v^{TS}$$

This study adopt the second approach because the multi-stage formulations in this study involved stochastic left hand side constraints, where the first approach could be infeasible due to too many variables are fixed from the deterministic problem. To avoid confusion, let EEV^{TS} be the expecion of expected value problem solution in two-stage stochastic programming. The relative value of two-stage stochastic programming (*RVSS*) and multi-stage stochastic programming (*RVMS*) are also defined as follows [122]:

$$RVSS \ = (v^{TS} - EEV^{TS})/EEV^{TS}$$

$$RVMS = (v^{MS} \ - v^{TS})/v^{TS}$$

However, the lower bound of *RVSS* and *RVMS* have more practical significance since both the two-stage and multi-stage models are hard to get the optimal solution. Let v_f^{TS} and v_f^{MS}be the objective value of the best feasible solution of two-stage and multi-stage models we could get in a reasonable comupational efforts, respectively. Let v_r^{TS} be the objective value of a relaxation of two-stage model, we could easily find out:

$$RVSS \ \geq (v_f^{TS} - EEV^{TS})/EEV^{TS}$$

$$RVMS \geq (v_f^{MS} \ - v_r^{TS})/v_r^{TS}$$

The numerical results and interpretations for measures of information are detailed discussed in the following results analysis sections.

5.4.4 Results analysis of deterministic model

The deterministic model yields to a total profit of $27494, which will be used as the objective value of *EV* solution. The seed with highest yield and highest plant population is selected by the model in the deterministic case. This is because under the average total water limits and precipitation amount, suitable irrigation decision could lead to no water stress. The NWO model under same conditions shows a total profit of $27137, which is almost the same the deterministic results. However, the deterministic model is oversimplified by using the mean of random variables to make decisions. A nature concern would be what will happen when there is water shortage and deficit irrigation therefore is needed? For each scenario, assuming we have prefect information before making decisions, the wait-and-see decisions could be found. The basic statistics of objective values for these wait-and-see decisions are summarized in Table 5.2. The average objective value of WS solutions is \$16790. These WS decisions are not implementable, however the WS solutions are the upper bound of profits under stochastic environment.s The significant profits drop from EV to WS indicates that the EV solution underestimate the effect of stochatic environments. Because in the EV solution, the first stage decision is made by ignoring the uncertainties. If we apply this EV solution in the stochastic environment, the objective value (profits) of EEV^{TS} ends up to be \$12127. EEV^{TS} decisions are easy to get and they are implenmentable. However, the performance is not good, as shown in Table 5.2.

Table 5.2 Basic statistics for WS and EEV objective values (Dollars)

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
WS	0	7220	18350	16790	26980	38440
EEV	-20120	1570	14940	12127	22200	33850

There is an information gap of \$4663 between WS and EEV^{TS} solutions, or the WS solution is 38.44% higher than the EEV^{TS} solution. This gap indicates applying stochastic programming may help to gain more value of information and lead to better decisions.

5.4.5 Results analysis of two-stage stochastic programming model

Before we go to the multi-stage stochastic programming, the two-stage stochastic programming is first investgated to calculate the $RVSS$ and verify the benefits of stochastic programming. Two-stage stochastic programming is a sepcial case of stochastic programming, which has a much shorter decision process. In the two-stage stochastic programming, the first stage still makes ($t = 0$) the pre-season decisions including the corn seed type selection and plant population selection. At the beginning of second stage ($t = 1$), realization of corn market price and seasonal irrigation water limits become available. The second stage decisions are how much irrigation water should be put in the field for the next eight weeks. These second stage decisions are made at the beginning of second stage. Note that the precipitation amount for the next eight weeks is not available when you make the second stage decisions, but these preicipitation information will be used to evaluate the objective values. The precisely decision process has the form:

decision of seed type and plant population→ observation of corn market price and seasonal irrigation water limits→ decision of irrigation for next eight weeks

Constraint (b15) should be changed to Constraint (c15) to reflect the change of decision process. Note that the two-stage stochastic programming is a special case of multi-stage stochastic programming, where decision maker has to make irrigation decision at a earlier time period. For maximization problem, the optimal solutions to the multi-stage problem will have a profit no less than the optimal solution to the two-stage problem because the multi-stage formulation's solution can adapt to information as it comes in. In other words, more stages allow more recourse and will yield to better (at least no worse) solutions.

$$y_{tw} = y_{tw'}, \ \forall w, w' \text{for which } \xi_{[t=2]}^{w} = \xi_{[t=2]}^{w'}, \forall t, \forall u \qquad \text{(c15)}$$

The objective value of two stage stochastic programming (v_f^{TS}) is \$13367, which yields a VSS of \$1239 and a $RVSS$ of 10.22%. These results could be interpreted as a 10.22% profit increase could be achieved by taking corn price and total water limits uncertainties into consideration when making the preseason decision of seed types seletion and plant populations seletion. Note that the uncertainties of precipitation are ignored in the two-stage decision process. The *EVPI* is \$3423 which also indicates that having additional information could potentially increase profit.

The same procedure of scenario generation and model solving are conducted ten times. The objective values of two-stage *RP* and *EEV* are summarized in Figure 5.2. The

v_f^{TS} for each time range from \$13114 to \$13933. These relative small ranges indicate that the scenario generation process is in-sample stability.

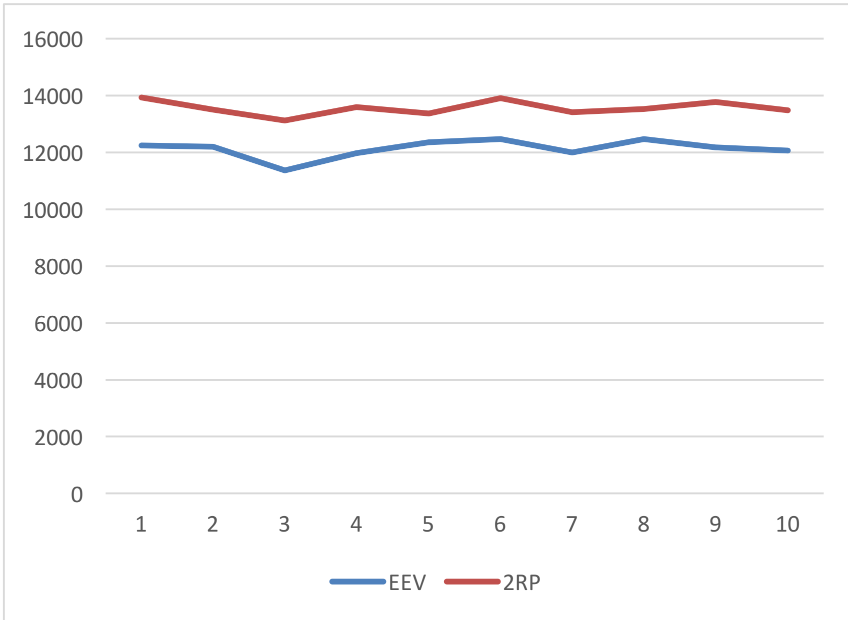


Figure 5.2 The objective values of two-stage RP and EEV for ten runs (Dollars)

5.4.6 Results analysis of multi-stage stochastic programming model

The objective value of multi-stage stochastic programming (v_f^{MS}) is \$15116, which yields a VMS of \$1749and a $RVMS$ of 13.08%. These results could be interpreted as a 13.08% profit increase could be achieved by taking precipitation uncertainties into consideration and use multi-stage decision process when making the preseason decision of seed types seletion and plant populations seletion. Weak out-of-sample stability test shows that two objective value of multi-stage stochastic programming by swifching the optimal decision is \$15116 and \$15304, respectively. This results indicate our model has out-of-sample stability.

Table 5.3 summaries the profit, decisions, and cost for different models. In the stochastic programming results, more conservative first stage decisions are made such as select high drought resistance seed. These decisions preform more robust in the stochastic environment. However, all models perfer high plant population, which indicates that the benefit of increasing yields is more significant than the drawback of increasing water demands for high plant population. It is worth noting that this effect might only hold when there is sufficient water (precipitation plus irrigation). Low plant population is still recommended at water-limited sites with no irrigation system.

Table 5.3 Comparison among different models (Dollars)

Model	Deterministic	Two-stage SP	Multi-stage SP
Total profit	27494	13367	15116
Sales of corn	113400	97157	99456
Production cost	62872	60018	60018
Irrigation cost	22891	23797	24278
Seed selection	high yield	high drought tolerance	high drought tolerance
Plant population	high	high	high

Although only the first stage decisions are implementable and all the later stage decisions are scenario based, it is still meanful to compare the average irrigation amount decisions for each model. Figure 5.3 summaries the average irrigation amount from each model for each week. As shown in Figure 5.3, the irrigation decisions in deterministic model are very progressive since it assumes the precipitation is deterministic and known.

The irrigation decisions for two-stage stochastic programming and multi-stage stochastic programming share the same pattern but the irrigation decisions for two-stage stochastic programming is more conservative. This is because little precipitation information are available for two-stage stochastic programming. The multi-stage stochastic programming can make recourse irrigation decisions based on the precipitation information at that point.

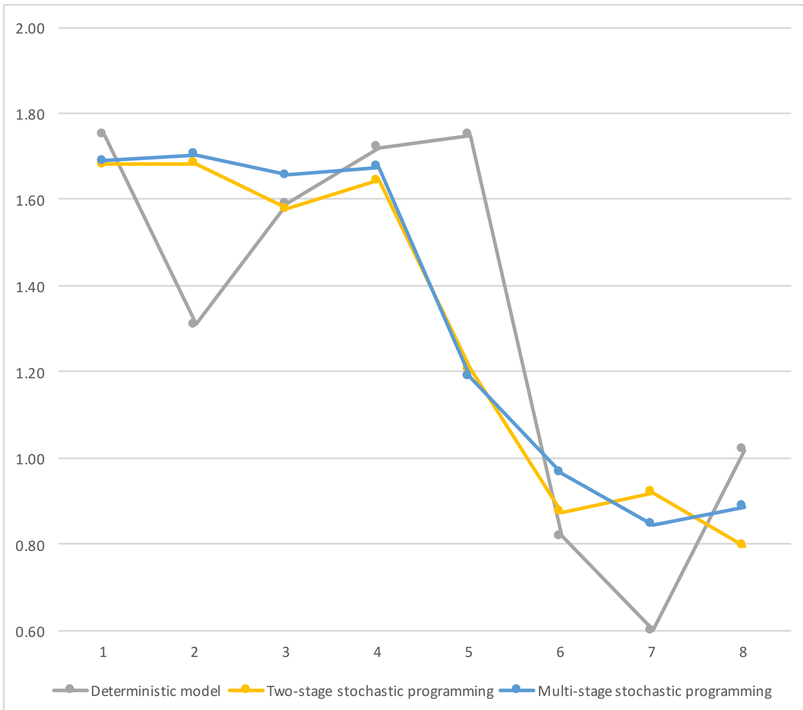


Figure 5.3 Comparison of weekly irrigation among different models (acre inches)

Note that the information releasing process is the same for deterministic model, two-stage stochastic programming, and multi-stage stochastic programming. The main difference among these model is the decision-making process. The deterministic model makes all decisions all at once, the two-stage stochastic programming separates the decision-making process into two stages, and the multi-stage stochastic programming makes sequence of decisions according to the stages. The case study results show that by delaying the decision-making process and considering more information (uncertainties), we could have higher profit by choosing better first-stage decisions.

5.5 Conclusion

In this study, a multi-stage stochastic programming model for farmland management under uncertainties is proposed. The first stage decisions include the pre-season decisions of seed types selection and plant populations selection, while the later stages determine when to irrigate and how much water should be put in the field during the corn flowering and grain filling stages. The uncertainties under investigated include corn price, irrigation water limits, and precipitation amount. Their distributions are carefully defined based on detailed derivation process. SAA method is used to generate scenarios.

The case study is based on a farm in Nebraska to illustrate and validate the optimization model. The numerical results show that a 10.22% profit increase could be achieved by taking corn price and total water limits uncertainties into consideration, and an additional 13.08% profit increase could be achieved by also taking precipitation uncertainties into consideration. These results indicate stochastic programming is a promising way for farmland management under uncertainties and can increase farmers' income significantly.

Our study is subject to a number of limitations. Firstly, the numerical results reported in the case study is the best feasible solution in a reasonable computational time. More efficient algorithm and heuristic solutions need to be investigated. Secondly, the case study only illustrates the model to a center pivot sprinkler systems with almost homogeneous soil features in Nebraska. Other irrigation systems and location could be also investigated. Thirdly, we consider three sources of independent uncertainties and more uncertainty factors can be considered. Last but not least, this model focus on a single year profit maximization problem. The evaluation of installation new irrigation system in a multi-year horizon is another interesting research problem. We shall address these limitations in our future research.

CHAPTER 6 SUMMARY AND DISCUSSION

This dissertation consists of four papers, and aims to contribute to the decision making methodology under uncertainties for renewable energy and precision agriculture. The contributions, limitations, and future works are discussed in this chapter.

The first paper provides a mathematical programming framework with a two-stage stochastic programming approach to deal with the uncertainties in the biofuel industry. The first stage makes capital investment decisions including the locations and capacities of facilities while the second stage determines the biomass and biofuels flow. This decision model focuses on dealing with uncertainties in a supply chain and can be easily adapted to deal with other uncertainties and be applied to other supply chain design problems. The optimization model also provides managerial suggestions for decision makers on the capital investment and logistic decisions in a stochastic environment. This study is subject to a number of limitations. Firstly, we assume the sustainability factor and farmers' participation are the same for each county. However, these factors may vary based on the land characteristics and agricultural management practices. Additional constraints such as water use constraints can be included to better describe biomass availability. Secondly, we assume the transportation cost within counties is negligible, which could impact the supply chain design and decision making. Thirdly, we consider three sources of uncertainties and more uncertainty factors can be considered. Last but not least, only one set of scenarios is generated in this paper; more scenarios could be generated to test the stability of the stochastic results. We shall address these limitations in our future research.

In the second paper, a new TEA method considering supply chain configurations has been introduced. The motivation of this proposed TEA method is to introduce supply chain design into traditional TEA to achieve a more comprehensive analysis and realistic economic assessment results. The proposed approach is illustrated with a case study to compare two competitive pathways of biofuel prodection in Iowa. The results indicate that biomass gasification pathway has better economic performance than hybrid fast pyrolysis and bio-oil gasification pathway under current technology status. Hybrid fast pyrolysis and bio-oil gasification pathway is more suitable for a decentralized supply chain structure while biomass gasification pathway is more suitable for a single centralized facility supply chain structure. As for the second paper, future study could achieve a more precise estimate by modeling the diversity of parameters such as scaling factors and labor costs. It should be noted that as a general framework, other biofuel production pathways could be evaluated considering supply chain configurations using the same procedures.

In the third paper, a farm-level precision farmland management problem for pre-season seed type selection and irrigation water management is introduced. A mixed integer linear program is proposed and variations of the basic model on different implementation conditions have been discussed. A case study based on a farmland in California has been conducted to demonstrate and validate the model. The model gives quantitative analysis for the farmland management. Special irrigation patterns are considered and the results show that the farmer's annual net profit could be significantly increased by applying the precision farming management decisions tools based on the proposed model. This model can also serve as a decision making and risk analysis tool for farmers facing seasonal irrigation water limits and extreme drought conditions. There are some future research directions based on the third paper as well. On one hand, the case study presented to illustrate and validate this optimization model only considers a certain piece of land. However, the shapes of farmland could affect the agricultural machinery paths and the homogeneous features of the soil could affect the shapes and sizes of decision units. Motivated by finite element analysis, models that allow different shapes and sizes of decision units in a piece of land. On the other hand, The proposed model could be used to evaluate other crops as well; and the interaction among plants such as plant population, leaf cover, and water competition could be stressed in future research.

In the fourth paper, a multi-stage stochastic programming model is formulated for farmland management under uncertainties. Precipitation amount, along with other uncertainties such as crop market price and irrigation water limits are investigated in the model. Optimal solutions for pre-season decisions and irrigation scheduling are given. The case study results indicate stochastic programming is a promising way for farmland management under uncertainties and can increase farmers' income significantly. This model contributes not only to the precision agriculture but also to protect water resources. There are some future research directions motivated by the fourth paper. Firstly, more efficient algorithm and heuristic solutions need to be investigated to find optimal solution in a reasonable computational time. Secondly, the case study only illustrates the model by a center pivot sprinkler systems with homogeneous soil features in Nebraska. Other irrigation systems, soil, and location could be also investigated. Thirdly, we consider three sources of independent uncertainties and more uncertainty factors can be considered. Last but not least, this model focus on a single year profit maximization problem. The evaluation of installation new irrigation system in a multi-year horizon is another interesting research problem. We shall address these limitations in future research.

BIBLIOGRAPHY

1. Timilsina, G.R. and A. Shrestha, *How much hope should we have for biofuels?* Energy, 2011. **36**(4): p. 2055-2069.

2. Carriquiry, M.A., X. Du, and G.R. Timilsina, *Second generation biofuels: economics and policies.* Energy Policy, 2011. **39**(7): p. 4222-4234.

3. McBratney, A., et al., *Future directions of precision agriculture.* Precision Agriculture, 2005. **6**(1): p. 7-23.

4. Zhang, N., M. Wang, and N. Wang, *Precision agriculture—a worldwide overview.* Computers and electronics in agriculture, 2002. **36**(2): p. 113-132.

5. Bongiovanni, R. and J. Lowenberg-DeBoer, *Precision agriculture and sustainability.* Precision agriculture, 2004. **5**(4): p. 359-387.

6. Schnepf, R., *Renewable Fuel Standard (RFS): overview and issues*. 2011: Diane Publishing.

7. Sethi, L.N., et al., *Optimal crop planning and conjunctive use of water resources in a coastal river basin.* Water resources management, 2002. **16**(2): p. 145-169.

8. Brown, T.R., *A techno-economic review of thermochemical cellulosic biofuel pathways.* Bioresource technology, 2015. **178**: p. 166-176.

9. Swanson, R.M., et al., *Techno-economic analysis of biomass-to-liquids production based on gasification.* Fuel, 2010. **89**: p. S11-S19.

10. Zhang, Y., et al., *Techno-economic analysis of two bio-oil upgrading pathways.* Chemical Engineering Journal, 2013. **225**: p. 895-904.

11. Li, Q., Y. Zhang, and G. Hu, *Techno-economic analysis of advanced biofuel production based on bio-oil gasification.* Bioresource technology, 2015. **191**: p. 88-96.

12. Sarker, R. and M. Quaddus, *Modelling a nationwide crop planning problem using a multiple criteria decision making tool.* Computers & Industrial Engineering, 2002. **42**(2): p. 541-553.

13. Tao, F., M. Yokozawa, and Z. Zhang, *Modelling the impacts of weather and climate variability on crop productivity over a large area: a new process-based model development, optimization, and uncertainties analysis.* agricultural and forest meteorology, 2009. **149**(5): p. 831-850.

14. Seppelt, R. and A. Voinov, *Optimization methodology for land use patterns—evaluation based on multiscale habitat pattern comparison.* Ecological Modelling, 2003. **168**(3): p. 217-231.

15. Shah, N., *Process industry supply chains: Advances and challenges.* Computers & Chemical Engineering, 2005. **29**(6): p. 1225-1235.

16. Ekşioğlu, S.D., et al., *Analyzing the design and management of biomass-to-biorefinery supply chain.* Computers & Industrial Engineering, 2009. **57**(4): p. 1342-1352.

17. Li, Y., T. Brown, and G. Hu, *Optimization model for a thermochemical biofuels supply network design.* Journal of Energy Engineering, 2014.

18. Li, Q. and G. Hu, *Techno-economic analysis of biofuel production considering logistic configurations.* Bioresource technology, 2016. **206**: p. 195-203.

19. Schnepf, R. and B.D. Yacobucci. *Renewable fuel standard (RFS): overview and issues*. in *CRS Report for Congress*. 2010.

20. Brown, R.C. and T.R. Brown, *Biorenewable resources: engineering new products from agriculture*. 2013: John Wiley & Sons.

21. Van Rossum, G., S.R. Kersten, and W.P. van Swaaij, *Catalytic and noncatalytic gasification of pyrolysis oil.* Industrial & engineering chemistry research, 2007. **46**(12): p. 3959-3967.

22. Megaritis, A., et al., *Pyrolysis and gasification in a bench-scale high-pressure fluidized-bed reactor.* Energy & fuels, 1998. **12**(1): p. 144-151.

23. Pollard, A., M. Rover, and R. Brown, *Characterization of bio-oil recovered as stage fractions with unique chemical and physical properties.* Journal of Analytical and Applied Pyrolysis, 2012. **93**: p. 129-138.

24. Badger, P.C. and P. Fransham, *Use of mobile fast pyrolysis plants to densify biomass and reduce biomass handling costs—A preliminary assessment.* Biomass and bioenergy, 2006. **30**(4): p. 321-325.

25. López-González, D., et al., *Gasification of lignocellulosic biomass char obtained from pyrolysis: Kinetic and evolved gas analyses.* Energy, 2014. **71**: p. 456-467.

26. Rogers, J. and J.G. Brammer, *Analysis of transport costs for energy crops for use in biomass pyrolysis plant networks.* Biomass and Bioenergy, 2009. **33**(10): p. 1367-1375.

27. An, H., W.E. Wilhelm, and S.W. Searcy, *Biofuel and petroleum-based fuel supply chain research: a literature review.* Biomass and Bioenergy, 2011. **35**(9): p. 3763-3774.

28. Nixon, J.D., et al., *Supply chain optimisation of pyrolysis plant deployment using goal programming.* Energy, 2014. **68**: p. 262-271.

29. Kim, J., M.J. Realff, and J.H. Lee, *Optimal design and global sensitivity analysis of biomass supply chain networks for biofuels under uncertainty.* Computers & Chemical Engineering, 2011. **35**(9): p. 1738-1751.

30. Marvin, W.A., et al., *Economic optimization of a lignocellulosic biomass-to-ethanol supply chain.* Chemical Engineering Science, 2012. **67**(1): p. 68-79.

31. Osmani, A. and J. Zhang, *Stochastic optimization of a multi-feedstock lignocellulosic-based bioethanol supply chain under multiple uncertainties.* Energy, 2013. **59**: p. 157-172.

32. Birge, J.R. and F.V. Louveaux, *Introduction to stochastic programming*. 1997: Springer.

33. Klingenfeld, D. and H. Kennedy, *Corn stover as a bioenergy feedstock: identifying and overcoming barriers for corn stover harvest, storage, and transport*. 2008: John F. Kennedy School of Government.

34. Wetterlund, E., et al., *Optimal localisation of biofuel production on a European scale.* Energy, 2012. **41**(1): p. 462-472.

35. Høyland, K. and S.W. Wallace, *Generating scenario trees for multistage decision problems.* Management Science, 2001. **47**(2): p. 295-307.

36. Tyndall, J.C., E.J. Berg, and J.P. Colletti, *Corn stover as a biofuel feedstock in Iowa's bio-economy: an Iowa farmer survey.* Biomass and Bioenergy, 2011. **35**(4): p. 1485-1495.

37. Wilcke, W. and G. Wyatt, *Grain Storage Tips.* Twin Cities, MN, The University of Minnesota Extension Service, the University of Minnesota, 2002.

38. Downing, M., et al., *US Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry*. 2011, Oak Ridge National Laboratory (ORNL).

39. USDA / NASS Quickstats. http://quickstats.nass.usda.gov/. *Accessed May 2013*.

40. Montross, M., et al. *Economics of collection and transportation of corn stover*. in *ASAE Paper 036081 presented at the Annual International Meeting of the American Society of Agricultural Engineers, Las Vegas, NV*. 2003.

41. Schechinger, T.M. and J. Hettenhaus, *Corn stover harvesting: Grower, custom operator, and processor issues and answers—report on corn stover harvest experiences in Iowa and Wisconsin for the 1997–98 and 1998–99 crop years.* ORNL/SUB-0404500008274-01. NTIS, Springfield, VA, 2004.

42. Lindstrom, M., *Effects of residue harvesting on water runoff, soil erosion and nutrient loss.* Agriculture, ecosystems & environment, 1986. **16**(2): p. 103-112.

43. Papendick, R.I. and W. Moldenhauer, *Crop residue management to reduce erosion and improve soil quality: Northwest.* Conservation research report, 1995.

44. Andrews, S.S., *Crop residue removal for biomass energy production: Effects on soils and recommendations.* White paper, USDA Natural Resources Conservation Service, 2006.

45. Graham, R.L., et al., *Current and potential US corn stover supplies.* Agronomy Journal, 2007. **99**(1): p. 1-11.

46. EIA, http://www.eia.gov/state/data.cfm?sid=IA - Consumption. *Accessed May 2013.*

47. EIA, http://www.eia.gov/oil_gas/petroleum/data_publications/wrgp/mogas_history.html. *Accessed May 2013*

48. U.S. Census Bureau, http://www.census.gov/population/metro/data/metrodef.html *Accessed Feb 2013.*

49. CBO, *Energy use in freight transportation. Congressional Budget Office (CBO).* 1982.

50. Searcy, E., et al., *The relative cost of biomass energy transport*, in *Applied Biochemistry and Biotecnology*. 2007, Springer. p. 639-652.

51. BTS. *Average Freight Revenue per Ton-mile,* http://www.bts.gov/publications/national_transportation_statistics/html/table_03_21.html.

52. Raffelt, K., et al., *The BTL2 process of biomass utilization entrained-flow gasification of pyrolyzed biomass slurries.* Applied Biochemistry and Biotechnology, 2006. **129**(1-3): p. 153-164.

53. Wright, M.M., et al., *Techno-economic analysis of biomass fast pyrolysis to transportation fuels.* Fuel, 2010. **89**: p. S2-S10.

54. Anex, R.P., et al., *Techno-economic comparison of biomass-to-transportation fuels via pyrolysis, gasification, and biochemical pathways.* Fuel, 2010. **89**: p. S29-S35.

55. Wright, M.M., R.C. Brown, and A.A. Boateng, *Distributed processing of biomass to bio‐oil for subsequent production of Fischer‐Tropsch liquids.* Biofuels, bioproducts and biorefining, 2008. **2**(3): p. 229-238.

56. Kazemzadeh, N. and G. Hu, *Optimization models for biorefinery supply chain network design under uncertainty.* Journal of Renewable and Sustainable Energy, 2013. **5**(5): p. 053125.

57. Wooley, R., et al., *Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis current and futuristic scenarios*. 1999, DTIC Document.

58. Grant, D., et al., *Feasibility of a producer-owned ground-straw feedstock supply system for bioethanol and other products.* Idaho: INL. 115p, 2006.

59. Papapostolou, C., E. Kondili, and J.K. Kaldellis, *Development and implementation of an optimisation model for biofuels supply chain.* Energy, 2011. **36**(10): p. 6019-6026.

60. Wang, S., et al., *Biogasoline production from the co-cracking of the distilled fraction of bio-oil and ethanol.* Energy & Fuels, 2013. **28**(1): p. 115-122.

61. Wang, S., et al., *Mechanism research on cellulose pyrolysis by Py-GC/MS and subsequent density functional theory studies.* Bioresource technology, 2012. **104**: p. 722-728.

62. Wang, S., et al., *Pyrolysis behaviors of four lignin polymers isolated from the same pine wood.* Bioresource technology, 2015. **182**: p. 120-127.

63. Wang, S., et al., *Degradation mechanism of monosaccharides and xylan under pyrolytic conditions with theoretic modeling on the energy profiles.* Bioresource technology, 2013. **143**: p. 378-383.

64. Zhang, Y., et al., *Comparative techno-economic analysis of biohydrogen production via bio-oil gasification and bio-oil reforming.* biomass and bioenergy, 2013. **51**: p. 99-108.

65. Harland, C.M., *Supply chain management, purchasing and supply management, logistics, vertical integration, materials management and supply chain dynamics.* Blackwell Encyclopedic dictionary of operations management. UK: Blackwell, 1996. **15**.

66. Manganaro, J.L. and A. Lawal, *Economics of thermochemical conversion of crop residue to liquid transportation fuel.* Energy & fuels, 2012. **26**(4): p. 2442-2453.

67. Zhang, Y. and M.M. Wright, *Product selection and supply chain optimization for fast pyrolysis and biorefinery system.* Industrial & Engineering Chemistry Research, 2014. **53**(51): p. 19987-19999.

68. Li, Q. and G. Hu, *Supply chain design under uncertainty for advanced biofuel production based on bio-oil gasification.* Energy, 2014. **74**: p. 576-584.

69. Zhang, L. and G. Hu, *Supply chain design and operational planning models for biomass to drop-in fuel production.* Biomass and Bioenergy, 2013. **58**: p. 238-250.

70. Sullivan, A., *Economics: Principles in action.* 2003.

71. Wright, M. and R.C. Brown, *Establishing the optimal sizes of different kinds of biorefineries.* Biofuels, Bioproducts and Biorefining, 2007. **1**(3): p. 191-200.

72. EIA, *Annual Energy Outlook*. 2014, US Department of Energy Washington DC.

73. Bridgwater, A.V., *Review of fast pyrolysis of biomass and product upgrading.* Biomass and bioenergy, 2012. **38**: p. 68-94.

74. Thilakaratne, R., et al., *Mild catalytic pyrolysis of biomass for production of transportation fuels: a techno-economic analysis.* Green Chem., 2014. **16**(2): p. 627-636.

75. Babcock, B., *Renewable identification numbers and the renewable fuels standard: how they support corn and soybean prices.* Iowa Ag Review, 2009. **15**(1): p. 7-9.

76. Argus. *Argus White Paper: Argus RINs prices*. 2013 [cited 2015; Available from: https://www.argusmedia.com/Methodology-and-Reference/Key-Prices/~/media/BEC7DB8BA70545928C0F883DC3A107D4.ashx.

77. Brown, P.D., T.A. Cochrane, and T.D. Krom, *Optimal on-farm irrigation scheduling with a seasonal water limit using simulated annealing.* Agricultural Water Management, 2010. **97**(6): p. 892-900.

78. Fereres, E. and M.A. Soriano, *Deficit irrigation for reducing agricultural water use.* Journal of experimental botany, 2007. **58**(2): p. 147-159.

79. Georgiou, P. and D. Papamichail, *Optimization model of an irrigation reservoir for water allocation and crop planning under various weather conditions.* Irrigation Science, 2008. **26**(6): p. 487-504.

80. Ganji, A., et al., *A new stochastic optimization model for deficit irrigation.* Irrigation Science, 2006. **25**(1): p. 63-73.

81. Singh, A., *Irrigation planning and management through optimization modelling.* Water resources management, 2014. **28**(1): p. 1-14.

82. Wardlaw, R. and K. Bhaktikul, *Application of genetic algorithms for irrigation water scheduling.* Irrigation and Drainage, 2004. **53**(4): p. 397-414.

83. Nagesh Kumar, D., K.S. Raju, and B. Ashok, *Optimal reservoir operation for irrigation of multiple crops using genetic algorithms.* Journal of Irrigation and Drainage Engineering, 2006. **132**(2): p. 123-129.

84. Gorantiwar, S. and I. Smout, *Multilevel approach for optimizing land and water resources and irrigation deliveries for tertiary units in large irrigation schemes. II: Application.* Journal of irrigation and drainage engineering, 2005. **131**(3): p. 264-272.

85. Yamout, G. and M. El-Fadel, *An optimization approach for multi-sectoral water supply management in the Greater Beirut Area.* Water Resources Management, 2005. **19**(6): p. 791-812.

86. Barton, B. and S.E. Clark, *Water & Climate Risks Facing US Corn Production: How Companies & Investors Can Cultivate Sustainability.* A Ceres Report. Boston, MA, 2014.

87. Eswaran, H., et al., *Soil classification: a global desk reference*. 2002: CRC Press.

88. Saxton, K. and W. Rawls, *Soil water characteristic estimates by texture and organic matter for hydrologic solutions.* Soil science society of America Journal, 2006. **70**(5): p. 1569-1578.

89. Campbell, G.S., *Soil physics with BASIC: transport models for soil-plant systems*. Vol. 14. 1985: Elsevier.

90. Saxton, K., et al., *Estimating generalized soil-water characteristics from texture.* Soil Science Society of America Journal, 1986. **50**(4): p. 1031-1036.

91. Wright, S., K. Klonsky, D. Stewart and others, *University of California Cooperative Extension 2015 Sample Costs Produce Field Corn on Mineral Soils in the Sacramento Valley. 2015.* Department of Agricultural and Resource Economics, UC Davis: Davis, CA, 2015.

92. Plastina, A., *Estimated Costs of Crop Production in Iowa, 2015.* Iowa State University Extension & Outreach, 2015.

93. Schwankl, L. and A. Fulton. *Corn ET Estimates*. 2015 [cited 2015 Feb.]; Available from: http://ucanr.edu/sites/Drought/files/167003.pdf.

94. Papendick, R.I. and W. Moldenhauer, *Crop residue management to reduce erosion and improve soil quality: Northwest.* Conservation research report (USA), 1995.

95. Edwards, W., *Estimating a value for corn stover.* Iowa State University Extension, Ag Decision Maker Document A1-70, 2014.

96. Vekshin, A. *California Water Prices Soar for Farmers as Drought Grows*. 2015 [cited 2016 Feb 2016]; Available from: http://www.bloomberg.com/news/articles/2014-07-24/california-water-prices-soar-for-farmers-as-drought-grows.

97. Itoh, T., H. Ishii, and T. Nanseki, *A model of crop planning under uncertainty in agricultural management.* International Journal of Production Economics, 2003. **81**: p. 555-558.

98. Spekken, M. and S. de Bruin, *Optimized routing on agricultural fields by minimizing maneuvering and servicing time.* Precision agriculture, 2013. **14**(2): p. 224-244.

99. Clark, J.S., et al., *Drought cycles and landscape responses to past aridity on prairies of the northern Great Plains, USA.* Ecology, 2002. **83**(3): p. 595-601.

100. Norwood, C.A. and T.J. Dumler, *Transition to dryland agriculture.* Agronomy Journal, 2002. **94**(2): p. 310-320.

101. Kranz, W., et al., *Irrigation Management for Corn. NebGuide G1850. Lincoln, Neb.: University of Nebraska–Lincoln*. 2008.

102. Grassini, P., H. Yang, and K.G. Cassman, *Limits to maize productivity in Western Corn-Belt: a simulation analysis for fully irrigated and rainfed conditions.* Agricultural and forest meteorology, 2009. **149**(8): p. 1254-1265.

103. Shaw, R.H., G. Sprague, and J. Dudley, *Climate requirement.* Corn and corn improvement., 1988(Ed. 3): p. 609-638.

104. Sudar, R.A., K.E. Saxton, and R. Spomer, *A predictive model of water stress in corn and soybeans.* Transactions of the ASAE, 1981. **24**(1): p. 97-0102.

105. Doerge, T., *Safely delaying the first irrigation of corn.* Crop Insights, 2008. **18**(7).

106. Doorenbos, J. and A. Kassam, *Yield response to water.* Irrigation and drainage paper, 1979. **33**: p. 257.

107. Kranz, W., et al., *Minimum center-pivot design capacities in Nebraska. NebGuide G1851. Lincoln, NE: University of Nebraska-Lincoln*. 2008.

108. Brown, P. *Flood vs. Pivot irrigation for Forage Crops: What are the Advantages and Disadvantages?* in *Proceedings, California Alfalfa & Forage Symposium and Western Seed Conference, San Diego, CA*. 2008.

109. Martin, D., et al. *Water Optimizer Single-Field Version: Instruction Manual*. 2010 [cited 2017 March 03]; Available from: http://agecon.unl.edu/79402687-fd71-43fa-914a-2c233956291e.pdf.

110. Blomvall, J. and A. Shapiro, *Solving multistage asset investment problems by the sample average approximation method.* Mathematical programming, 2006. **108**(2-3): p. 571-595.

111. Chepuri, K. and T. Homem-De-Mello, *Solving the vehicle routing problem with stochastic demands using the cross-entropy method.* Annals of Operations Research, 2005. **134**(1): p. 153-181.

112. Høyland, K., M. Kaut, and S.W. Wallace, *A heuristic for moment-matching scenario generation.* Computational optimization and applications, 2003. **24**(2-3): p. 169-185.

113. King, A.J. and S.W. Wallace, *Modeling with stochastic programming*. 2012: Springer Science & Business Media.

114. Arnold, C. and W. Elliot, *CLIGEN weather generator predictions of seasonal wet and dry spells in Uganda.* Transactions of the ASAE, 1996. **39**(3): p. 969-972.

115. Hanson, C., et al., *Microcomputer program for daily weather simulation in the contiguous United States.* ARS (USA), 1994.

116. Husak, G.J., J. Michaelsen, and C. Funk, *Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications.* International Journal of Climatology, 2007. **27**(7): p. 935-944.

117. Alam, A.J., M.S. Rahman, and A. Sadaat, *Markov Chain Analysis of Weekly Rainfall Data for Predicting Agricultural Drought*, in *Computational Intelligence Techniques in Earth and Environmental Sciences*. 2014, Springer. p. 109-128.

118. Heitsch, H. and W. Römisch, *Scenario tree modeling for multistage stochastic programs.* Mathematical Programming, 2009. **118**(2): p. 371-406.

119. Madansky, A., *Inequalities for stochastic linear programming problems.* Management science, 1960. **6**(2): p. 197-204.

120. Birge, J. and F. Louveaux, *Introduction to Stochastic Programming. Series in Operations Research and Financial Engineering*. 1997, Springer.

121. Escudero, L.F., et al., *The value of the stochastic solution in multistage problems.* Top, 2007. **15**(1): p. 48-64.

122. Huang, K. and S. Ahmed, *The value of multistage stochastic programming in capacity planning under uncertainty.* Operations Research, 2009. **57**(4): p. 893-904.